Functional Analysis

Swayam Chube

January 7, 2025

Contents

§1 Preliminaries on Topological Vector Spaces

Lemma 1.1 (Riesz Lemma). Let X be a normed linear space and $Y \subsetneq X$ a proper closed subspace. Then, for every $0 < \alpha < 1$, there is an $x \in X \setminus Y$ such that $\|x\| = 1$ and $\operatorname{dist}(x, Y) > \alpha$.

§2 Completeness Arguments

Theorem 2.1 (Rudin, Exercise 4.26). Let X and Y be Banach spaces. The set of all surjective bounded linear operators in $\mathscr{B}(X, Y)$ forms an open subset.

Proof. Let $T : X \to Y$ be a surjective linear operator. By the open mapping theorem, there is an $r > 0$ such that $B_Y(0, 2r) \subseteq T\left(B_X(0,1)\right)$. If $0 \neq y \in Y$, then $\frac{ry}{\|y\|} \in B_Y(0, 2r)$, consequently, there is an $x' \in X$ with $\|x'\| < 1$ and $Tx' = \frac{ry}{\|y\|}$, thus, $x = \frac{\|y\|}{r}x'$ maps to y under T. Note that $\|x\| < \frac{\|y\|}{r}$. For the sake of brevity, let $t = 1/r$.

Let $\delta = \frac{1}{2t'} > 0$ and $S \in \mathscr{B}(X, Y)$ such that $\|T - S\| < \delta$. We shall show that S is surjective, for which, it would suffice to show that the image of S contains the unit ball of Y. Indeed, let $y_0 \in Y$ with $\|y_0\| \leqslant 1$. Choose an $x_0 \in X$ such that $\|x_0\| < t$ and $Tx_0 = y_0$. Setting $y_1 = y_0 - Sx_0$, we have

$$\|y_1\| = \|(T - S)x_0\| \leqslant \delta t.$$

Again, choose $x_1 \in X$ such that $Tx_1 = y_1$ and $\|x_1\| < t\|y_1\| = \delta t^2$. Setting $y_2 = y_1 - Sx_1$, we have

$$\|y_2\| = \|(T - S)x_1\| \leqslant \delta^2 t^2$$

and so on. We have thus constructed two sequences $(x_n)_{n \geqslant 0}$ and $(y_n)_{n \geqslant 0}$ such that

- $Tx_n = y_n$,
- $y_{n+1} = y_n - Sx_n$ for $n \geqslant 0$, and
- $\|x_n\| < \delta^n t^{n+1}$ and $\|y_n\| \leqslant \delta^n t^n$.

Let $x = \sum_{n=0}^{\infty} x_n$, which converges since $\sum_{n=0}^{\infty} \|x_n\|$ does. Hence,

$$Sx = \lim_{n \to \infty} \sum_{i=0}^{n} Sx_i = \sum_{i=0}^{\infty} y_i - y_{i+1} = y_0,$$

thereby completing the proof. ∎

§3 THE HAHN-BANACH THEOREMS

LEMMA 3.1 (DOMINATED EXTENSION THEOREM). Let X be a real vector space with a subspace M. Suppose $p : X \to \mathbb{R}$ satisfies

$$p(x+y) \leqslant p(x) + p(y) \quad \text{and} \quad p(tx) = tp(x) \quad \forall x, y \in M,\ \forall t \geqslant 0.$$

Let $f : X \to \mathbb{R}$ be a linear functional such that $f(x) \leqslant p(x)$ for all $x \in M$. Then, there is a linear functional $\Lambda : X \to \mathbb{R}$ such that $\Lambda x = f(x)$ for all $x \in M$ and

$$-p(-x) \leqslant \Lambda x \leqslant p(x) \quad \forall x \in X.$$

Proof. If $M = X$, then there is nothing to prove. Suppose now that M is a proper subspace of X and choose $x_1 \in X \setminus M$. For $x, y \in M$, we have

$$f(x) + f(y) = f(x+y) \leqslant p(x+y) \leqslant p(x - x_1) + p(y + x_1),$$

and hence,

$$f(x) - p(x - x_1) \leqslant -f(y) + p(y + x_1) \quad \forall x, y \in M.$$

Let α denote the supremum of the left hand side in the above inequality as x ranges over M. Note that α is finite as the left hand side is always bounded above by $p(x_1)$. Let $M_1 = M + \mathbb{R}x_1$ and define $f_1 : M_1 \to \mathbb{R}$ by

$$f_1(m + \lambda x_1) = f(m) + \lambda\alpha;$$

in particular, $f_1(x_1) = \alpha$. Note that for $\lambda \neq 0$,

$$\begin{aligned} f_1(m + \lambda x_1) &= |\lambda| f_1(|\lambda|^{-1} m + \operatorname{sgn}(\lambda) x_1) \\ &= |\lambda| f(|\lambda|^{-1} m) + \lambda\alpha \\ &\leqslant |\lambda| \Big(p(|\lambda|^{-1} m + \operatorname{sgn}(\lambda) x_1) - \operatorname{sgn}(\lambda)\alpha \Big) \\ &= p(m + \lambda x_1). \end{aligned}$$

This furnishes an extension $f_1 : M_1 \to \mathbb{R}$ such that $f_1(y) \leqslant p(y)$ for all $y \in M_1$. One can then extend this, using Zorn's Lemma, to $\Lambda : X \to \mathbb{R}$ such that $\Lambda x \leqslant p(x)$ for all $x \in X$. We then have

$$-p(-x) \leqslant -\Lambda(-x) = \Lambda x \leqslant p(x),$$

thereby commpleting the proof. ■

THEOREM 3.2 (HAHN-BANACH EXTENSION THEOREM). Let M be a subspace of a vector space (real or complex) X, p a semi-norm on X, and f a linear functional on M such that $|f(x)| \leqslant p(x)$ for all $x \in M$. Then f extends to a linear functional Λ on X satisfying $|\Lambda x| \leqslant p(x)$ for all $x \in X$.

Proof. Suppose first that the field of scalars is $\mathbb{R}$. Due to the preceding lemma, f can be extended to $\Lambda : X \to \mathbb{R}$ satisfying

$$-p(x) = -p(-x) \leqslant \Lambda x \leqslant p(x) \quad \forall x \in X,$$

that is, $|\Lambda x| \leqslant p(x)$.

Next, suppose the field of scalars is $\mathbb{C}$. Let $u = \Re f$. Due to the first part of the proof, u can be extended to a linear functional $U : X \to \mathbb{R}$ satisfying $|Ux| \leqslant p(x)$ for all $x \in X$. Define $\Lambda : X \to \mathbb{C}$ by

$$\Lambda x = u(x) - iu(ix) \quad \forall x \in X.$$

We contend that Λ is the desired functional. Let $x \in X$ and choose an $\alpha \in \mathbb{C}$ with $|\alpha| = 1$ such that $\alpha \Lambda x = |\Lambda x|$. Hence,

$$|\Lambda x| = \alpha \Lambda x = \underbrace{\Lambda(\alpha x) = U(\alpha x)}_{\text{because LHS } \in \mathbb{R}_{\geqslant 0}} \leqslant p(\alpha x) = p(x).$$

This completes the proof. ■

COROLLARY. Let X be a normed linear space and M a subspace of X. Suppose $f : M \to \mathbb{K}$ is a bounded linear functional, then there exists a bounded linear functional $\Lambda : X \to \mathbb{K}$ extending f. Further, $\|f\| = \|\Lambda\|$

Proof. Invoke the preceding result with $p(x) = \|f\|\|x\|$. I ■

THEOREM 3.3 (HAHN-BANACH SEPARATION THEOREM). Suppose A and B are disjoint convex subsets of a topological vector space X.

(a) If A is open, there exist $\Lambda \in X^*$ and $\gamma \in \mathbb{R}$ such that

$$\Re\Lambda x < \gamma \leqslant \Re\Lambda y \quad \forall x \in A,\ y \in B.$$

(b) If A is compact, B is closed, and X is locally convex, there exist $\Lambda \in X^*$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that

$$\Re\Lambda x < \gamma_1 < \gamma_2 < \Re\Lambda y \quad \forall x \in A,\ y \in B.$$

Proof. We first prove this theorem when the scalar field is assumed to be $\mathbb{R}$.

(a) Fix points $a_0 \in A$, $b_0 \in B$. Set $x_0 = b_0 - a_0$ and $C = A - B + x_0$. Then, C is a convex neighborhood of 0 in X, and thus, admits a Minkowski functional, $p : X \to \mathbb{R}$ which is subadditive and $p(tx) = tp(x)$ for all $t \geqslant 0$. Further, since $A \cap B = \varnothing$, $x_0 \notin C$, whence $p(x_0) \geqslant 1$.

Define a linear functional $f : \mathbb{R}x_0 \to \mathbb{R}$ by $f(\lambda x_0) = \lambda$ and using the Dominated Extension Theorem, extend this to a functional $\Lambda : X \to \mathbb{R}$ such that

$$-p(-x) \leqslant \Lambda x \leqslant p(x) \quad \forall x \in X.$$

Let $D = C \cap (-C)$, which is a symmetric convex neighborhood of the origin. For any $x \in D$, it is easy to see that $p(x) \leqslant 1$, whence

$$-1 \leqslant -p(-x) \leqslant \Lambda x \leqslant p(x) \leqslant 1,$$

and hence, Λ is a continuous linear functional.

Now, for $a \in A$ and $b \in B$,

$$\Lambda a - \Lambda b = \Lambda(a-b) = \Lambda(a-b+x_0) - 1 \leqslant p(a-b+x_0) - 1 < 0,$$

since $a - b + x_0 \in C$. Hence, $\Lambda a < \Lambda b$ for every $a \in A$ and $b \in B$. Finally, since $\Lambda(A)$ and $\Lambda(B)$ are disjoint convex subsets of $\mathbb{R}$, both must be intervals with the former to the left of the latter. Further, since the former is an open subset of $\mathbb{R}$, we immediately obtain the desired conclusion.

(b) There is a convex, balanced neighborhood V of the origin in X such that $(A+V) \cap (B+V) = \emptyset$. Set $C = A + V$, which is a convex open subset of X, disjoint from B. Due to part (a), there is a linear functional Λ such that $\Lambda(C)$ is to the left of $\Lambda(B)$ and $\Lambda(A)$ sits as a compact interval inside $\Lambda(C)$. The conclusion now is immediate.

We now suppose that the field of scalars is $\mathbb{C}$; whence X is also a topological $\mathbb{R}$-vector space. In both parts (a) and (b), we were able to obtain an $\mathbb{R}$-linear functional, continuous on X when viewed as a $\mathbb{R}$-TVS and separating the two sets as desired. Define the $\mathbb{C}$-linear functional $\Lambda x = u(x) - iu(ix)$ and note that this has the desired separation properties too. ■

COROLLARY. If X is an LCTVS, then X^* separates points on X.

Proof. Let $p, q \in X$. Use Theorem 3.3 (b) with $A = \{p\}$ and $B = \{q\}$. ■

THEOREM 3.4. Let M be a proper closed subspace of a locally convex topological vector space, and $x_0 \in X \setminus M$. There exists a linear functional $\Lambda \in X^*$ such that $\Lambda x_0 = 1$ and $\Lambda x = 0$ for all $x \in M$.

Proof. Using Theorem 3.3(b) with $A = \{x_0\}$ and $B = M$, there is a $\Lambda \in X^*$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that

$$\Re \Lambda x_0 < \gamma_1 < \gamma_2 < \Re \Lambda y \quad \forall y \in M.$$

Since $\Lambda(0) = 0$ and $0 \in M$, we must have that $\Lambda x_0 \neq 0$. Further, since $\lambda y \in M$ for every $\lambda \in \mathbb{K}$, the only way $\Re(\lambda \Lambda y) > \gamma_2$ for every $\lambda \in \mathbb{K}$ is if Λ vanishes on M. Dividing Λ by Λx_0, we have our desired conclusion. ■

COROLLARY. Let X be an LCTVS and $M \subseteq X$ a subspace. Suppose $f : M \to \mathbb{K}$ is a continuous linear functional, then there is a $\Lambda \in X^*$ such that $\Lambda|_M = f$.

Proof. ■

§4 Weak and Weak* Topologies

Lemma 4.1. Let X be a $\mathbb{K}$-vector space and $\Lambda_1, \ldots, \Lambda_n, \Lambda$ be linear functionals on X and set

$$N = \{x \in X \colon \Lambda_i x = 0, \ \forall 1 \leqslant i \leqslant n\}.$$

The following are equivalent:

(a) There are scalars $\alpha_1, \ldots, \alpha_n \in \mathbb{K}$ such that

$$\Lambda = \alpha_1 \Lambda_1 + \cdots + \alpha_n \Lambda_n.$$

(b) There exists $0 < \gamma < \infty$ such that

$$|\Lambda x| \leqslant \gamma \max_{1 \leqslant i \leqslant n} |\Lambda_i x| \quad \forall x \in X.$$

(c) $\Lambda x = 0$ for every $x \in N$.

Proof. $(a) \implies (b) \implies (c)$ is trivial. It remains to show that $(c) \implies (a)$. Consider the map $\Phi : X \to \mathbb{K}^n$ given by

$$\Phi(x) = (\Lambda_1 x, \ldots, \Lambda_n x)$$

and let $Y \subseteq \mathbb{K}^n$ be its image. Define $\Psi : Y \to \mathbb{K}$ by

$$\Psi(\Phi(x)) = \Lambda x.$$

That this is well-defined follows from the fact that $N \subseteq \ker \Lambda$. Since we are in a finite-dimensional space, the map Ψ can be extended to a linear map $\Psi : \mathbb{K}^n \to \mathbb{K}$, which must be of the form

$$(y_1, \ldots, y_n) \mapsto \alpha_1 y_1 + \cdots + \alpha_n y_n.$$

It then follows that $\Lambda = \alpha_1 \Lambda_1 + \cdots + \alpha_n \Lambda_n$. ■

Definition 4.2. Let X be a set and

$$\mathscr{F} = \{f : X \to Y_f\}$$

a collection of functions. The *$\mathscr{F}$-topology* on X is defined to be the coarsest topology such that every $f \in \mathscr{F}$ is continuous.

The set $\mathscr{F}$ is said to *separate points* if for each pair $p \neq q$ in X, there is an $f \in \mathscr{F}$ such that $f(p) \neq f(q)$.

Remark 4.3. The $\mathscr{F}$-topology is more explicitly the topology generated by

$$\{f^{-1}(U) \colon U \subseteq Y_f \text{ is open}, \ f \in \mathscr{F}\}.$$

Proposition 4.4. If $\mathscr{F}$ is a separating family of functions on a space X, and each Y_f is Hausdorff, then the $\mathscr{F}$-topology on X is Hausdorff.

Proof. Let $p \neq q$ be points in X and choose $f \in \mathscr{F}$ such that $f(p) \neq f(q)$. Then, there are disjoint neighborhoods U and V of $f(p)$ and $f(q)$ respectively in Y_f. Since each f is continuous, $f^{-1}(U)$ and $f^{-1}(V)$ are disjoint neighborhoods of p and q in the $\mathscr{F}$-topology. ■

PROPOSITION 4.5. If X is a compact topological space and $\mathscr{F}$ is a countable family of continuous separating real-valued functions on X, then X is metrizable.

Proof. Let $\mathscr{F} = \{f_n \colon n \geqslant 1\}$. We may suppose without loss of generality that $\|f\|_\infty \leqslant 1$ for each $f \in \mathscr{F}$. It is not hard to check that the function $d : X \times X \to \mathbb{R}$ given by

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} |f_n(x) - f_n(y)|$$

is a metric inducing the topology on X. ■

THEOREM 4.6. Let X be a $\mathbb{K}$-vector space and X' a vector space of linear functionals on X that separates points. The X'-topology τ' on X makes it a locally convex topological vector space whose dual is X'.

Proof. Due to Proposition 4.4, τ' is Hausdorff. Note that the topology is generated by the set

$$\{\Lambda^{-1}(U) \colon \Lambda \in X',\ U \subseteq \mathbb{K} \text{ is open}\}.$$

Hence, a base for the topology is given by finite intersections of elements of the above form. Thus, is generated by intersections of the form

$$\Lambda_1^{-1}(U_1) \cap \cdots \cap \Lambda_n^{-1}(U_n),$$

where $U_1, \ldots, U_n \subseteq \mathbb{K}$ are open sets. It immediately follows that this base is translation invariant whence, the entire topology is translation invariant. A local base at 0 is given by open sets of the above form, such that $0 \in U_i$ for $1 \leqslant i \leqslant n$. We can further refine this local base by choosing open sets of the form

$$V(\Lambda_1, \ldots, \Lambda_n; \varepsilon_1, \ldots, \varepsilon_n) = \{x \in X \colon |\Lambda_i x| \leqslant \varepsilon_i,\ 1 \leqslant i \leqslant n\}.$$

Further, from this description, it is not hard to see that αV is a basic open set whenever $\alpha > 0$ and V a basic open set.

Now that we have established a local base for τ', we show that (X, τ') is indeed a topological vector space. That τ' is locally convex immediately follows from the above description of a local base. Next, we show that addition is continuous, for which it suffices to show continuity at $(0, 0) \in X \times X$. Let U be a neighborhood of 0 in X, then U contains a basic open set V of the above form. Since $\frac{1}{2}V + \frac{1}{2}V \subseteq V$, we see that addition is continuous.

To see that scalar multiplication is continuous, let $x \in X$, $\alpha \in \mathbb{K}$ and $x + V$ a neighborhood of x. We may suppose that V is a basic open set of the above form. Since V is absorbing, there is an $s > 0$ such that $x \in sV$. Choose r sufficiently small so that $r(r + s) + r|\alpha| < 1$. Then, if $y \in x + rV$, and $|\beta - \alpha| < r$,

$$\beta y - \alpha x = (\beta - \alpha)y + \alpha(y - x) \in r(r + s)V + |\alpha| rV \subseteq V,$$

since $y \in (r+s)V$. Hence, scalar multiplication is continuous and (X, τ') is a locally convex topological vector space.

Finally, let Λ be a continuous linear functional on X and consider a basic open set $V(\Lambda_1, \ldots, \Lambda_n, \varepsilon_1, \ldots, \varepsilon_n)$ such that $|\Lambda x| < 1$ on V. Thus, there is a $\gamma > 0$ such that

$$|\Lambda x| \leqslant \gamma \max_{1 \leqslant i \leqslant n} |\Lambda_i x|$$

whence, Λ is a linear combination of the Λ_i. ∎

DEFINITION 4.7. Let X be a topological vector space whose dual X^* separates points on X (this is true in particular for locally convex TVSs). Then the X^*-topology on X is called the *weak topology* and is denoted by (X, τ_w) or X_w.

Obviously the weak topology is coarser than the original topology. A set $E \subseteq X$ is said to be *weakly bounded* if it is bounded in the weak topology. Similarly, a sequence (x_n) is said to be *weakly convergent* to x if it converges in the weak topology. Since the weak topology is Hausdorff, the limit of any weakly convergent sequence is unique.

PROPOSITION 4.8. Let X be a topological vector space on which X^* separates points. Then

(a) X_w is a locally convex topological vector space.

(b) A set $E \subseteq X$ is weakly bounded if and only if every $\Lambda \in X^*$ is bounded on E.

(c) A sequence (x_n) is weakly convergent to x if and only if $\Lambda x_n \to \Lambda x$ for every $\Lambda \in X^*$.

Proof. All three assertions are trivial. ∎

PROPOSITION 4.9. Let X be a locally convex topological vector space and $E \subseteq X$ a convex subset. Then the weak closure $\overline{E}_w$ is the same as the original closure $\overline{E}$.

Proof. Since the weak topology is coarser than the original topology, $\overline{E} \subseteq \overline{E}_w$. Now, let $x_0 \in X \setminus \overline{E}$. Due to the Hahn-Banach Separation Theorem, there is an $\Lambda \in X^*$ and $\gamma_1, \gamma_2 \in \mathbb{R}$ such that

$$\Re \Lambda x_0 < \gamma_1 < \gamma_2 < \Re \Lambda y \quad \forall y \in \overline{E} \supseteq E.$$

Thus, there is a weak neighborhood of x_0 not intersecting E, consequently, $x_0 \notin \overline{E}_w$. This completes the proof. ∎

THEOREM 4.10. Suppose X is an infinite-dimensional normed linear space. Then the weak topology on X is not metrizable.

Proof. We shall show that the weak topology (X, w) is not first-countable whence the conclusion would follow. Suppose not, then there is a local base $\{U_n\}$ at 0. For each $n \geqslant 1$, there is a finite subset $F_n \subseteq X^*$ and $\varepsilon_n > 0$ such that

$$V_n = \{x \in X \colon |f(x)| < \varepsilon_n, \ \forall f \in F_n\}.$$

We contend that

$$X^* = \bigcup_{n \geqslant 1} \operatorname{span}(F_n).$$

Indeed, let $g \in X^*$ and

$$U = \{x \in X \colon |g(x)| < 1\}.$$

There is an index $n \geqslant 1$ such that $V_n \subseteq U$. Now, if x is in $\bigcap_{f \in F_n} \ker f$, then so is λx for every $\lambda \in \mathbb{K}$, consequently, $\lambda x \in V_n$ and hence, $|\lambda||g(x)| < 1$ for every $\lambda \in \mathbb{K}$. This forces $g(x) = 0$, that is,

$$\bigcap_{f \in F_n} \ker f \subseteq \ker g,$$

which, in light of Lemma 4.1 gives $g \in \operatorname{span}(F_n)$, proving our claim.

It follows that X^* has at most countable dimension and since X is infinite-dimensional, so is X^*, but this is absurd, since X^* is a Banach space. ∎

Definition 4.11. Let X be a topological vector space and X^*. The evaluation functionals induced by X form a separating vector space of functionals. The X-topology induced on X^* by these functionals is called the *weak* topology*.

Theorem 4.12 (Banach-Alaoglu). Let X be a topological vector space and V a neighborhood of 0. The *polar* of V:

$$K = \{\Lambda \in X^* \colon |\Lambda x| \leqslant 1,\ \forall x \in V\} \subseteq X^*$$

is weak*-compact.

Proof. Since V is a neighborhood of the origin, it is absorbing and hence, for each $x \in X$, there is $\gamma(x) > 0$ such that $x \in \gamma(x)V$. For $x \in V$, choose $\gamma(x) \leqslant 1$. Let D_x denote the compact set

$$D_x = \{z \in \mathbb{K} \colon |z| \leqslant \gamma(x)\}\,, \tag{1}$$

and

$$P = \prod_{x \in X} D_x,$$

which is compact due to Tychonoff's Theorem. Further, for each $\Lambda \in K$ and $x \in X$, since $x/\gamma(x) \in V$, we have $|\Lambda x| \leqslant |\gamma(x)|$, consequently, the element $(\Lambda x)_{x \in X}$ is an element of P. Thus, we can identify K with a subset of P. Henceforth, we shall denote elements of P as functions $f : X \to \mathbb{K}$. We shall show that:

(i) the subspace topology K inherits from P and the weak*-topology on K are the same,

(ii) with respect to the subspace topology, K is closed in P;

whence it follows that K is compact.

Let $\Lambda_0 \in K$ and consider a basic open set in the weak*-topology centered at Λ_0 of the form

$$W = \{\Lambda \in X^* \colon |\Lambda x_i - \Lambda_0 x_i| < \varepsilon,\ 1 \leqslant i \leqslant n\}\,.$$

In the product topology on P, the following set is open

$$V = \{f \in P \colon |f(x_i) - \Lambda_0 x_i| < \varepsilon,\ 1 \leqslant i \leqslant n\}\,.$$

It is not hard to see that $W \cap K = V \cap K$. This shows that the subspace topology induced on K by the product topology is finer than that induced by the weak*-topology.

On the other hand, choose any open set in the product topology in P intersecting K and choose an element Λ_0 in the intersection. The aforementioned open set contains one of the form V as above and since $W \cap K = V \cap K$, we see that the weak*-topology is finer than the subspace topology. This shows that the two topologies are the same.

Finally, we must show that K is closed in P. Let $f_0 \in \overline{K}$, $x, y \in X$ and $\alpha, \beta \in \mathbb{K}$. We contend that $f_0(\alpha x + \beta y) = \alpha f_0(x) + \beta f_0(y)$. Let $\varepsilon > 0$ and

$$V = \{f \in P \colon |f(z) - f_0(z)| < \varepsilon,\ z \in \{x, y, \alpha x + \beta y\}\}.$$

There is some $f \in K \cap V$. Then,

$$\begin{aligned}
&|f_0(\alpha x + \beta y) - \alpha f(x) - \beta f(y)| \leqslant \\
&|f_0(\alpha x + \beta y) - f(\alpha x + \beta y)| + |\alpha f(x) - \alpha f_0(x)| + |\beta f(y) - \beta f_0(y)| \\
&\leqslant (|\alpha| + |\beta| + 1)\varepsilon.
\end{aligned}$$

Since the above inequality holds for all $\varepsilon > 0$, we have that f_0 is linear. Further, by construction, f_0 is bounded by 1 on V, since $\gamma(x) \leqslant 1$ for all $x \in V$ and hence, $f_0 \in X^*$. It follows that $f_0 \in K$ and hence, K is closed in P, thereby completing the proof. ■

PROPOSITION 4.13 (RUDIN, EXERCISE 3.11). Let X be an infinite dimensional Fréchet space. Then X^* with the weak*-topology is of the first category in itself.

Proof. Let $V_n = B(0, 1/n) \subseteq X$ and let K_n denote their respective polars, that is

$$K_n = \{\Lambda \in X^* \colon |\Lambda x| \leqslant 1,\ \forall x \in V_n\}.$$

First, we claim that $X^* = \bigcup_{n=1}^{\infty} K_n$. Indeed, for any $\Lambda \in X^*$, note that the open set $\Lambda^{-1}(B_{\mathbb{K}}(0,1))$ contains some V_n and hence, $\Lambda \in K_n$.

It remains to now show that these have empty interior. Indeed, suppose K_N has nonempty interior for some $N \in \mathbb{N}$. Since K_N is convex, symmetric, so is its interior. Thus, we have that 0 lies in the interior of K_N. As a result, there is an $\varepsilon > 0$ and $x_1, \ldots, x_n \in X$ such that

$$W = \{\Lambda \in X^* \colon |\Lambda x_i| < \varepsilon,\ 1 \leqslant i \leqslant n\} \subseteq K_N.$$

Since K_N is compact, it is bounded and hence, so is W. But since X^* is infinite-dimensional too, so is $\bigcap_{i=1}^{n} \ker \mathrm{ev}_{x_i} \subseteq W$ which is contained in a bounded set, whence, must be the trivial subspace.

Next, for any $x \in X$, note that

$$\bigcap_{i=1}^{n} \ker \mathrm{ev}_{x_i} = \{0\} \subseteq \ker \mathrm{ev}_x,$$

thus x is a linear combination of the x_i's, that is, X is finite-dimensional, a contradiction. This completes the proof. ■

§§ The Krein-Milman Theorem

DEFINITION 4.14. A subset E of a topological vector space X is said to be *totally bounded* if to every neighborhood V of 0 in X corresponds a finite set F such that $E \subseteq F + V$.

REMARK 4.15. Note that we can require that $F \subseteq E$. Indeed, let V be a neighborhood of 0 and choose a neighborhood W of 0 such that $W + W \subseteq V$. There is a finite set $F \subseteq X$ such that $E \subseteq F + W$. For each $f \in F$ such that $(f + W) \cap E \neq \emptyset$, choose some e in the intersection. For any $w \in W$, we have $f + w - e = (f - e) + w \in W + W \subseteq V$. Hence, $f + W \subseteq e + V$. The collection of all such e's, say $\widetilde{F}$ is such that $E \subseteq \widetilde{F} + W$

THEOREM 4.16. (a) If $A_1, \ldots, A_n$ are compact convex sets in a topological vector space X, then $\mathrm{co}(A_1 \cup \cdots \cup A_n)$ is compact.

(b) If X is an LCTVS and $E \subseteq X$ is totally bounded, then $\mathrm{co}(E)$ is totally bounded.

(c) If X is a Fréchet space and $K \subseteq X$ is compact, then $\overline{\mathrm{co}}(X)$ is compact.

Proof. (a) Let

$$\Delta = \{(s_1, \ldots, s_n) \in \mathbb{R}^n \colon s_1 + \cdots + s_n = 1,\ s_i \geqslant 0\ \forall 1 \leqslant i \leqslant n\}.$$

Let $A = A_1 \times \cdots \times A_n$ and define the map $f : \Delta \times A \to X$ by

$$f(s, a) = s_1 a_1 + \cdots + s_n a_n.$$

This is a continuous map since addition and scalar multiplication are continuous on X. Put $K = f(S \times A)$. Then, K is compact and is contained in $\mathrm{co}(A_1 \cup \cdots \cup A_n)$.

We shall show that $K = \mathrm{co}(A_1 \cup \cdots \cup A_n)$, for which is suffices to show that K is convex (since each A_i is contained in K). Indeed, let $\alpha, \beta > 0$ with $\alpha + \beta = 1$. Then, for $(s, a), (t, b) \in S \times A$, we have

$$\alpha \sum_{i=1}^{n} s_i a_i + \beta \sum_{i=1}^{n} t_i b_i = \sum_{i=1}^{n} (\alpha s_i + \beta t_i) \cdot \frac{\alpha s_i a_i + \beta t_i b_i}{\alpha s_i + \beta t_i} = f(u, c),$$

where $u = \alpha s + \beta t$ and

$$c_i = \frac{\alpha s_i a_i + \beta t_i b_i}{\alpha s_i + \beta t_i} \in A_i,$$

and we are done.

(b) Let U be a neighborhood of 0 in X and choose a convex, balanced neighborhood V of 0 in X such that $V + V \subseteq U$. There is a finite set $F \subseteq X$ such that $E \subseteq F + V$, whence $E \subseteq \mathrm{co}(F) + V$. Since the latter is convex, we have $\mathrm{co}(E) \subseteq \mathrm{co}(F) + V$.

Due to part (a), $\mathrm{co}(F)$ is compact. The collection $\{f + V \colon f \in \mathrm{co}(F)\}$ is an open cover of $\mathrm{co}(F)$ and hence, admits a finite subcover, $\mathrm{co}(F) \subseteq F_1 + V$ for some $F_1 \subseteq X$. Therefore,

$$\mathrm{co}(E) \subseteq F_1 + V + V \subseteq F_1 + U,$$

that is, $\mathrm{co}(E)$ is totally bounded.

(c) Due to part (b), $\operatorname{co}(K)$ is totally bounded. Thus, its closure is totally bounded and complete, whence compact. ■

LEMMA 4.17 (CARATHÉODORY). If $E \subseteq \mathbb{R}^n$ and $x \in \operatorname{co}(E)$, then x lies in the convex hull of of some subset of E which contains at most $n+1$ points.

Proof. We shall show that if $k > n$ and $x = \sum_{i=1}^{k+1} t_i x_i$ is a convex combination for some $x_i \in \mathbb{R}^n$, then x is a convex combination of some k of these vectors. This is enough to prove the statement of the theorem.

We may suppose without loss of generality that $t_i > 0$ for $1 \leqslant i \leqslant k+1$. Consider the linear map $\mathbb{R}^{k+1} \to \mathbb{R}^{n+1}$ given by

$$(a_1, \dots, a_{k+1}) \mapsto \left(\sum_{i=1}^{k+1} a_i x_i, \sum_{i=1}^{k+1} a_i \right).$$

The kernel of this map must be nontrivial and hence, there exists $(a_1, \dots, a_{k+1}) \in \mathbb{R}^{k+1}$ with some $a_i \neq 0$, so that $\sum_{i=1}^{k+1} a_i x_i = 0$ and $\sum_{i=1}^{k+1} a_i = 0$. Set

$$|\lambda| = \min_{1 \leqslant i \leqslant k+1} \frac{t_i}{|a_i|},$$

which is finite, since $a_i \neq 0$ for some $1 \leqslant i \leqslant k+1$. Choose the sign of λ so that $\lambda a_j = \lambda_j$ for some $1 \leqslant j \leqslant k+1$. Set $c_i = t_i - \lambda a_i \geqslant 0$. Then,

$$\sum_{i=1}^{k+1} c_i x_i = \sum_{i=1}^{k+1} t_i x_i - \lambda \sum_{i=1}^{k+1} a_i x_i = x,$$

and

$$\sum_{i=1}^{k+1} c_i = \sum_{i=1}^{k+1} t_i - \lambda \sum_{i=1}^{k+1} a_i = 1.$$

Note that $c_j = 0$ and hence, we have written x as a convex combination of some k of the x_i's. ■

PROPOSITION 4.18. If $K \subseteq \mathbb{R}^n$ is compact, then so is $\operatorname{co}(K)$.

Proof. Let

$$\Delta = \{(s_1, \dots, s_{n+1}) \in \mathbb{R}^{n+1} \colon s_1 + \cdots + s_{n+1} = 1,\ s_i \geqslant 0\ \forall 1 \leqslant i \leqslant n+1\}.$$

Due to Carathéodory's lemma, it follows that $x \in \operatorname{co}(K)$ if and only if x is a linear combination of some $n+1$ elements of K. Thus, the map $\Delta \times K^{n+1} \to \mathbb{R}^n$ given by

$$(t, x_1, \dots, x_{n+1}) \mapsto t_1 x_1 + \cdots + t_{n+1} x_{n+1}$$

is continuous and its image is $\operatorname{co}(K)$. This completes the proof. ■

DEFINITION 4.19. Let X be a $\mathbb{K}$-vector space and $K \subseteq X$. A non-empty set $S \subseteq K$ is called an *extreme set* of K if whenever $x, y \in K$, $0 < t < 1$ such that $(1-t)x + ty \in S$, then $x, y \in S$.

The *extreme points* of K are the extreme sets that are singletons. The set of all extreme points of K is denoted by $E(K)$.

LEMMA 4.20. Let X be a topological vector space on which X^* separates points. Suppose A, B are disjoint, nonempty, compact, convex sets in X. Then there exists $\Lambda \in X^*$ such that

$$\sup_{x \in A} \Re \Lambda x < \inf_{y \in B} \Re \Lambda y.$$

Proof. Topologize X with the weak topology, which is coarser than the original topology, and hence, A, B are compact. Now, use the Hahn-Banach separation theorem and the fact that $(X_w)^* = X^*$. ∎

THEOREM 4.21 (KREIN-MILMAN). Let X be a topological vector space on which X^* separates points. If $K \subseteq X$ is a nonempty compact convex set in X, then $K = \overline{\mathrm{co}}(E(K))$.

Proof. Let $\mathscr{P}$ denote the poset of all nonemtpy compact extreme sets of K ordered by inclusion. Note that $\mathscr{P}$ is nonempty, since $K \in \mathscr{P}$. We make the following two observations about $\mathscr{P}$:

(a) If $S \neq \varnothing$, is an intersection of elements of $\mathscr{P}$, then $S \in \mathscr{P}$.

(b) If $S \in \mathscr{P}$, $\Lambda \in X^*$ and $\mu = \max\limits_{x \in S} \Re \Lambda x$, then

$$S_\Lambda = \{x \in S \colon \Re \Lambda x = \mu\} \in \mathscr{P}.$$

Observation (a) is obvious. As for (b), first note that S_Λ is closed in S, and hence, in K, thus, is compact. Now, suppose $x, y \in K$ and $t > 0$ such that $tx + (1-t)y \in S_\Lambda \subseteq S$. Since S is an extreme set of K, $x, y \in S$, consequently, $\Re \Lambda x, \Re \Lambda y \leqslant \mu$ and

$$\mu = \Re \Lambda (tx + (1-t)y) \leqslant t\mu + (1-t)\mu = \mu,$$

whence $x, y \in S_\Lambda$, thereby proving (b).

Choose some $S \in \mathscr{P}$ and let $\mathscr{P}'$ be the sub-poset of all members of $\mathscr{P}$ that are contained in S. Let Ω be a maximal chain in $\mathscr{P}'$ and let M denote the intersection of all elements of Ω. Since Ω has the finite intersection property and all sets in Ω are compact, $M \neq \varnothing$ and is compact.

We contend that M is a singleton. Indeed, since $M_\Lambda \subseteq M$, due to the minimality of M, we must have that $M_\Lambda = M$ for every $\Lambda \in X^*$. That is, $\Re \Lambda (x - y) = 0$ for all $x, y \in M$ and $\Lambda \in X^*$. Since X^* separates points on X, we must have that $x - y = 0$, that is, M is a singleton.

We have therefore proved that $E(K) \cap S \neq \varnothing$ for every $S \in \mathscr{P}$. Now, since K is convex, $\overline{\mathrm{co}}(E(K)) \subseteq K$, consequently, the former is compact. Suppose now that there is some

$x_0 \in K \setminus \overline{\text{co}}(E(K))$. Applying the preceding lemma with $B = \{x_0\}$ and $A = \overline{\text{co}}(E(K))$, there is a $\Lambda \in X^*$ such that

$$\Re \Lambda x_0 > \sup_{y \in \overline{\text{co}}(E(K))} \Re \Lambda y.$$

Then, $K_\Lambda \in \mathscr{P}$ and is disjoint from $\overline{\text{co}}(E(K))$, a contradiction. Thus, $\overline{\text{co}}(E(K)) = K$, thereby completing the proof. ■

§5 Compact Operators

Definition 5.1. A linear map $T : X \to Y$ between Banach spaces is said to be *compact* if $T(U)$ is relatively compact in Y where U is the unit ball in X.

The following proposition is immediate from the equivalence of compactness and sequential compactness in metric spaces.

Proposition 5.2. *T is compact if and only if every bounded sequence (x_n) in X contains a subsequence (x_{n_k}) such that (Tx_{n_k}) converges in Y.*

Definition 5.3. The *spectrum* $\sigma(T)$ of an operator $T \in \mathscr{B}(X)$ is the set of all scalars λ such that $T - \lambda I$ is not invertible.

Theorem 5.4. *Let X and Y be Banach spaces.*

(a) If $T \in \mathscr{B}(X, Y)$ and $\dim \mathscr{R}(T) < \infty$, then T is compact.

(b) If $T \in \mathscr{B}(X, Y)$, T is compact, and $\mathscr{R}(T)$ is closed, then $\dim \mathscr{R}(T) < \infty$.

(c) The compact operators form a closed subspace of $\mathscr{B}(X, Y)$ in its norm-topology.

(d) If $T \in \mathscr{B}(X)$, T is compact, and $\lambda \neq 0$ is a scalar, then $\dim \mathscr{N}(T - \lambda I) < \infty$.

(e) If $\dim X = \infty$, $T \in \mathscr{B}(X)$, and T is compact, then $0 \in \sigma(T)$.

(f) If $S, T \in \mathscr{B}(X)$, and T is compact, then so are ST and TS.

Proof. (a) Let U denote the unit ball of X. Then $T(U)$ is a bounded subset of $\mathscr{R}(T)$ and since the latter is closed in Y, $\overline{T(U)}$ is a closed and bounded subset of $\mathscr{R}(T)$, consequently, is compact.

(b) Since $\mathscr{R}(T)$ is closed in Y, it is complete, i.e., a Banach space. Due to the open mapping theorem, $T(U)$ is open in $\mathscr{R}(T)$ with compact closure, whence $\mathscr{R}(T)$ is locally compact, and hence, finite dimensional.

(c) Let $T_n \to T$ in $\mathscr{B}(X, Y)$ where each T_n is a compact operator. We shall show that $T(U)$ is totally bounded in Y. Let $\varepsilon > 0$ and choose an N such that $\|T - T_N\| < \varepsilon/3$. Note that $T_N(U)$ is totally bounded in Y, and hence, there are $x_1, \dots, x_n \in U$ such that

$$T_N(U) \subseteq \bigcup_{i=1}^{n} B_Y(T_N x_i, \varepsilon/3).$$

Now, for any $y \in U$, there is an index $1 \leqslant i \leqslant n$ such that $T_N y \in B(T_N x_i, \varepsilon/3)$. As a result,

$$\|Ty - Tx_i\| \leqslant \|Ty - T_N y\| + \|T_N y - T_N x_i\| + \|T_N x_i - Tx_i\| < \varepsilon.$$

Hence,

$$T(U) \subseteq \bigcup_{i=1}^{n} B_Y(Tx_i, \varepsilon),$$

and the conclusion follows.

(d) Let $Y = \mathscr{N}(T - \lambda I)$. Then note that T acts on Y by $y \mapsto \lambda y$. Further, since T is compact and Y is closed in X, the restriction of T to Y is compact and hence, Y must be finite-dimensional.

(e) If $0 \notin \sigma(T)$, then T is invertible, whence $\mathscr{R}(T)$ is closed but $\dim \mathscr{R}(T) = \infty$, a contradiction.

(f) This follows from Proposition 5.2. ■

THEOREM 5.5. Suppose X and Y are Banach spaces and $T \in \mathscr{B}(X, Y)$. Then T is compact if and only if $T^* \in \mathscr{B}(Y^*, X^*)$ is compact.

Proof. Suppose first that T is compact and let $\{y_n^*\}$ be a sequence in the unit ball of Y^*. We shall show that $T^* y^* = y^* \circ T$ admits a convergent subsequence in X^*. Let $K = \overline{T(U)} \subseteq Y$, which, according to our assumption is compact in Y. Note that the collection $\{y_n^*\}$ is equicontinuous and pointwise bounded on K. Due to the Arzelá-Ascoli Theorem, there is a subsequence $\{y_{n_k}^*\}$ that converges uniformly on K.

We contend that $\{T^* y_{n_k}^*\}$ converges in the operator norm. Indeed, for any $x \in U$,

$$|(T^* y_{n_k}^*(x) - T^* y_{n_l}^*(x)| = |y_{n_k}^*(Tx) - y_{n_l}^*(Tx)|,$$

and since $Tx \in K$, the conclusion follows.

Conversely, suppose T^* is compact. Consider the natural isometric embeddings $\Phi_X : X \to X^{**}$ and $\Phi_Y : Y \to Y^{**}$, which fit into a commutative diagram

$$\begin{tikzcd} X \arrow[r, "X"] \arrow[d, "\Phi_X"'] & Y \arrow[d, "\Phi_Y"] \\ X^{**} \arrow[r, "T^{**}"'] & Y^{**}. \end{tikzcd} \tag{2}$$

Due to the first part of the proof, T^{**} is compact. Thus, $T^{**}(U^{**})$ is totally bounded in Y^{**}. Next, $\Phi_X(U)$ is contained in U^{**} and hence, $T^{**}\Phi_X(U) = \Phi_Y T(U)$ is totally bounded in Y^{**}. Since Φ_Y is an isometry, it follows that $T(U)$ is totally bounded in Y, thereby completing the proof. ■

DEFINITION 5.6. A closed subspace M of a topological vector space X is said to be *complemented* if there exists a closed subspace N of X such that

$$X = M + N \quad \text{and} \quad M \cap N = \{0\}.$$

In this case, X is said to be the *direct sum* of M and N, denoted by $X = M \oplus N$.

LEMMA 5.7. Let M be a closed subspace of a topological vector space X.

(a) If X is locally convex and $\dim M < \infty$, then M is complemented in X.

(b) If $\dim(X/M) < \infty$, then M is complemented in X.

Proof. (a) Let $\{e_1, \dots, e_n\}$ be a basis for M. Every $x \in M$ has a unique representation

$$x = \alpha_1(x)e_1 + \cdots + \alpha_n(x)e_n.$$

Note that $\alpha_i(e_j) = 0$ whenever $i \neq j$. Due to the Hahn-Banach Theorem, each α_i can be extended to a continuous linear functional on X. Let $N = \bigcap_{i=1}^{n} \mathscr{N}(\alpha_i)$. It is not hard to argue that $X = M \oplus N$.

(b) Let $\pi : X \to X/M$ be the quotient map, and let $\{e_1, \dots, e_n\}$ be a basis for X/M. Lift this to $\{x_1, \dots, x_n\}$ in X and let N be the vector subspace they span. Again, it is not hard to argue that $X = M \oplus N$. ■

THEOREM 5.8. Let X be a Banach space, $T \in \mathscr{B}(X)$ a compact operator, and $\lambda \neq 0$. Then $T - \lambda I$ has closed range.

Proof. Let $N = \mathscr{N}(T - \lambda I)$, which is a closed subspace of X. Due to Lemma 5.7, admits a complement, say M. Let $S : M \to X$ be given by $x \mapsto Tx - \lambda x$, which is a bounded linear operator. Since $\mathscr{R}(S) = \mathscr{R}(T - \lambda I)$, it suffices to show that the former is closed.

To this end, we first show that there is a constant $\beta > 0$ such that $\|Sx\| \geqslant \beta\|x\|$ for all $x \in M$, which is equivalent to

$$\beta = \inf_{\substack{\|x\|=1 \\ x \in M}} \|Sx\| > 0.$$

Suppose not. Then, there is a sequence $x_n \in M$ with $\|x_n\| = 1$, such that $Sx_n \to 0$ as $n \to \infty$. Since $T : X \to X$ is compact, its restriction to M is also compact, whence, there is a subsequence (x_{n_k}) such that $Tx_{n_k} \to x_0$ for some $x_0 \in X$. Replace x_n with this subsequence. Then, $Tx_n - \lambda x_n \to 0$ and hence, $\lambda x_n \to x_0$. As a result,

$$Sx_0 = \lim_{n \to \infty} S(\lambda x_n) = \lambda \lim_{n \to \infty} Sx_n = 0.$$

But since S is injective, $x_0 = 0$. This is absurd, since $\|x_0\| = \lim_{n \to \infty} \|\lambda x_n\| = |\lambda| > 0$. It follows that $\beta > 0$.

Finally, we show that $\mathscr{R}(S)$ is closed in X. Indeed, suppose $y \in \overline{\mathscr{R}(S)}$; then there is a sequence (x_n) in M such that $Sx_n \to y$, that is (Sx_n) is Cauchy. But since

$$\beta\|x_n - x_m\| \leqslant \|Sx_n - Sx_m\|,$$

so is (x_n). Hence, $x_n \to x_0$ for some $x_0 \in M$; and $Sx_0 = y$. This completes the proof. ■

THEOREM 5.9 (SPECTRUM OF A COMPACT OPERATOR). Let X be a Banach space and $T \in \mathscr{B}(X)$ a compact operator.

(a) Every $0 \neq \lambda \in \sigma(T)$ is an eigenvalue of T.

(b) For every $\lambda \neq 0$, the increasing chain of subspaces

$$\mathscr{N}(T-\lambda I) \subseteq \mathscr{N}((T-\lambda I)^2) \subseteq \cdots$$

eventually stabilizes. Further, a these subspaces are finite dimensional.

(c) For every $r > 0$, the set

$$\{\lambda \in \sigma(T) \colon |\lambda| > r\}$$

is finite.

(d) As a consequence, $\sigma(T)$ is countable and the only possible limit point of $\sigma(T)$ is 0.

Proof. Suppose $\dim X = \infty$, for if $\dim X < \infty$, then all the above statements are trivial as there are only finitely many eigenvalues.

(a) Suppose $0 \neq \lambda \in \sigma(T)$ is not an eigenvalue of T, then $T - \lambda I$ is injective, but not surjective, else, due to the open mapping theorem, it would be invertible. Define

$$Y_n = (T-\lambda I)^n(X).$$

Obviously, $Y_{n+1} \subseteq Y_n$ for all $n \geqslant 1$. Further, since the restriction of T to each of these subspaces is compact, due to Theorem 5.4 (d), each Y_n is infinite-dimensional and all inclusions are strict.

For each $n \geqslant 1$, using the Riesz Lemma, choose $y_n \in Y_n \setminus Y_{n+1}$ such that $\|y_n\| = 1$ and

$$\operatorname{dist}(y_n, Y_{n+1}) > \frac{1}{2}.$$

Since T is compact and (x_n) is bounded, the sequence (Tx_n) must admit a convergent subsequence. But for $n < m$, we have

$$\|Tx_n - Tx_m\| = \|(T-\lambda I)x_n + \lambda x_n - (T-\lambda I)x_m - \lambda x_m\|,$$

and since $(T-\lambda I)x_n - (T-\lambda I)x_m - \lambda x_m \in Y_{n+1}$, we conclude that $\|Tx_n - Tx_m\| > \lambda/2$, a contradiction.

(b) If λ is not an eigenvalue, then each $\mathscr{N}((T-\lambda I)^n)$ is the trivial subspace and there is nothing to prove. Suppose now that λ is an eigenvalue of T and set $Y_n = \mathscr{N}((T-\lambda I)^n)$. Obviously $Y_1 \subseteq Y_2 \subseteq \cdots$. Further, $(T-\lambda I)^n = S + (-\lambda)^n I$ where S is some compact operator and hence, $\dim Y_n < \infty$. Next, note that if $Y_n = Y_{n+1}$ for some $n \geqslant 1$, then $Y_n = Y_{n+1} = Y_{n+2} = \cdots$.

Suppose now that $Y_n \subsetneq Y_{n+1}$ for every $n \geqslant 1$. Again, using the Riesz Lemma, choose $y_{n+1} \in Y_{n+1} \setminus Y_n$ such that $\|y_{n+1}\| = 1$ and

$$\operatorname{dist}(y_{n+1}, Y_n) > \frac{1}{2}.$$

Again, since (y_n) is bounded and T is compact, the sequence (Ty_n) must admit a convergent subsequence. But for $2 \leqslant n < m$, we have

$$\|Ty_n - Ty_m\| = \|(T - \lambda I)y_n + \lambda y_n - (T - \lambda I)y_m - \lambda y_m\|,$$

and since $(T-\lambda I)y_n - (T-\lambda I)y_m + \lambda y_n \in Y_{m-1}$, it follows that $\|Ty_n - Tx_m\| > \lambda/2$, a contradiction.

(c) Suppose there is an $r > 0$ such that the set $\{\lambda \in \sigma(T) \colon |\lambda| > r\}$ is infinite. Choose a countable subset $\{\lambda_1, \lambda_2, \dots\}$ with corresponding eigenvectors $\{x_1, x_2, \dots\}$. Let $Y_n = \operatorname{span}\{x_1, \dots, x_n\}$; when then form a strictly increasing chain of closed subspaces.

First, we contend that for $n \geqslant 2$, $(T - \lambda_n I)(Y_n) \subseteq Y_{n-1}$. Indeed, any element of Y_n can be written uniquely as

$$Y_n \ni y = \alpha_1 x_1 + \cdots + \alpha_n x_n.$$

Then, $(T - \lambda_n I)y = \alpha_1(T - \lambda_n I)x_1 + \cdots + \alpha_{n-1}(T - \lambda_n I)x_{n-1}$. And for $1 \leqslant i \leqslant n-1$, we have

$$(T - \lambda_i I)(T - \lambda_n)x_i = (T - \lambda_n I)(T - \lambda_i I)x_i = 0,$$

whence $(T - \lambda_n)x_i \in Y_i$.

Next, using the Riesz Lemma, for $n \geqslant 2$, choose $y_n \in Y_n \setminus Y_{n-1}$ such that $\|y_n\| = 1$ and

$$\operatorname{dist}(y_n, Y_{n-1}) > \frac{1}{2}.$$

Since (y_n) is bounded and T is compact, the sequence (Ty_n) admits a convergent subsequence. But for $2 \leqslant n < m$, we have

$$\|Ty_n - Ty_m\| = \|(T - \lambda_n I)y_n + \lambda_n y_n - (T - \lambda_m I)y_m - \lambda_m y_m\|,$$

and since

$$(T - \lambda_n I)y_n + \lambda_n y_n - (T - \lambda_m I)y_m \in Y_{m-1},$$

we get that $\|Ty_n - Ty_m\| > |\lambda_m|/2 > r/2$, a contradiction.

(d) Note that

$$\sigma(T) = \{0\} \cup \bigcup_{n \geqslant 1} \left\{\lambda \in \sigma(T) \colon |\lambda| > \frac{1}{n}\right\},$$

and being a countable union of finite sets, $\sigma(T)$ is countable. Next, suppose $0 \neq \mu \in \mathbb{K}$ is a limit point of $\sigma(T)$. There is an $\varepsilon > 0$ such that $|\mu| > \varepsilon$. But since the set of eigenvalues in $\mathbb{K} \setminus \overline{B}(0, \varepsilon)$ is finite, μ cannot be their limit point. This completes the proof. ■

§§ Examples

THEOREM 5.10 (MINKOWSKI'S INTEGRAL INEQUALITY). Let $(X, \mathfrak{M}, \mu)$ and $(Y, \mathfrak{N}, \lambda)$ be positive measure spaces. If $f : X \times Y \to \mathbb{R}$ is non-negative and measurable with respect to the product measure, then for $1 \leqslant p < \infty$,

$$\left\{ \int_X \left(\int_Y f(x,y)\, d\lambda(y) \right)^p d\mu(x) \right\}^{\frac{1}{p}} \leqslant \int_Y \left(\int_X f(x,y)^p\, d\mu(x) \right)^{\frac{1}{p}} d\lambda(y)$$

Proof. Since $p = 1$ is just Fubini, we assume $p > 1$ and let q be the conjugate exponent to p. Let $H : X \to \mathbb{R}$ be defined as

$$H(x) = \int_Y f(x,y)\, d\lambda(y),$$

which is a measurable function on X due to Fubini. We now have the series of inequalities

$$\begin{aligned} \|H\|_p^p &= \int_X \int_Y f(x,y) H(x)^{p-1}\, d\lambda(y) d\mu(x) \\ &= \int_Y \int_X f(x,y) H(x)^{p-1}\, d\mu(x) d\lambda(y) \\ &\leqslant \int_Y \left(\int_X f(x,y)^p d\mu(x) \right)^{\frac{1}{p}} \left(\int_X H(x)^{pq-q} \right)^{\frac{1}{q}} d\lambda(y) \\ &= \int_Y \left(\int_X f(x,y)^p\, d\mu(x) \right)^{\frac{1}{p}} \|H\|_p^{\frac{p}{q}}\, d\lambda(y) \end{aligned}$$

and hence

$$\|H\|_p \leqslant \int_X \left(\int_X f(x,y)^p\, d\mu(x) \right)^{\frac{1}{p}} d\lambda(y),$$

thereby completing the proof. ∎

THEOREM 5.11. Let $1 < p < \infty$ and define the *Hardy operator* $H : L^p(0,\infty) \to L^p(0,\infty)$ as

$$Hf(x) = \frac{1}{x} \int_0^x f(t)\, dt.$$

Then, H is a non-compact operator with operator norm

$$\|H\| = \frac{p}{p-1}.$$

Proof. For operator norm, take $x^{-1/p}\chi_{[0,N]}$ and let $N \to \infty$. ∎

§6 REFLEXIVE SPACES

DEFINITION 6.1. A normed linear space X is said to be *reflexive* if the natural embedding $\Phi : X \to X^{**}$ is surjective.

PROPOSITION 6.2. Let X be a normed linear space. The natural embedding $\Phi : X \to X^{**}$ is a topological imbedding when X is given the weak topology and X^* is given the weak*-topology.

Proof. ■

THEOREM 6.3 (KAKUTANI). A Banach space X is reflexive if and only if its norm-closed unit ball is weakly compact.

Proof. Let B, B^{**} denote the norm-closed unit balls of X and X^{**} respectively. If X were reflexive, then the natural embedding $\Phi : X \to X^{**}$ is surjective. Due to the preceding result, Φ is a homeomorhpism when X is given the weak topology and X^{**} is given the weak*-topology. Since B^{**} is compact in the weak*-topology, and Φ is an isometry, we see that B must be compact in the weak topology.

Conversely, suppose B is compact in the weak topology. Again, due to the preceding proposition, $\Phi(B)$ is compact and convex in the weak*-topology and $\Phi(B) \subseteq B^{**}$. If X were not reflexive, then $\Phi(B) \subsetneq B^{**}$. Choose $x^{**} \in B^{**} \setminus \Phi(B)$. Due to the Hahn-Banach Separation Theorem, there is a linear functional $\Lambda : X^{**} \to \mathbb{K}$ that is continuous with respect to the weak*-topology on X^* and there are $\gamma_1, \gamma_2 \in \mathbb{R}$ such that

$$\Re\Lambda(x^{**}) < \gamma_1 < \gamma_2 < \Re\Lambda(y) \quad \forall y \in \Phi(B).$$

Note that there is some $0 \neq x^* \in X^*$ such that $\Lambda = \mathrm{ev}_{x^*}$, and hence,

$$\Re x^{**}(x^*) < \gamma_1 < \gamma_2 \leqslant \inf_{y \in \Phi(B)} \Re y(x^*) = \inf_{x \in B} \Re x^*(x).$$

The rightmost quantity is precisely $-\|x^*\|$. Thus $\Re x^{**}(x^*) < -\|x^*\|$, in particular, $|x^{**}(x^*)| > \|x^*\|$, whence $\|x^{**}\| > 1$, a contradiction, since we chose it inside B^{**}. This completes the proof. ■

COROLLARY. Every closed, bounded convex subset of a reflexive Banach space is weakly compact.

Proof. This follows from the fact that a convex closed subset of an LCTVS is also weakly closed. ■

§7 HILBERT SPACES

DEFINITION 7.1. An *inner product space* is a $\mathbb{K}$-vector space H together with a function $(\cdot,\cdot) : H \times H \to \mathbb{K}$ such that

(i) $(x, y) = \overline{(y, x)}$,

(ii) $(x + y, z) = (x, z) + (y, z)$,

(iii) $(\alpha x, y) = \alpha(x, y)$,

(iv) $(x, x) \geqslant 0$, and $(x, x) = 0$ if and only if $x = 0$, for all $x, y, z \in H$ and $\alpha \in \mathbb{K}$.

Obviously, $\|x\| := \sqrt{(x,x)}$ defines a norm on H. If H is complete with respect to this norm, then H is said to be a *Hilbert space*.

PROPOSITION 7.2. Let H be an inner product space and $x, y \in H$. Then,

$$|(x,y)| \leqslant \|x\|\|y\| \quad \text{and} \quad \|x+y\| \leqslant \|x\| + \|y\|.$$

Proof. For every $\lambda \in \mathbb{K}$, we have

$$0 \leqslant (x+\lambda y, x+\lambda y) = |\lambda|^2\|y\|^2 + \|x\|^2 + 2\Re(x, \lambda y).$$

For every $\alpha \in \mathbb{R}$, we can choose $\lambda \in \mathbb{K}$ such that $|\lambda| = |\alpha|$ and $\Re(x, \lambda y) = \alpha|(x,y)|$. Thus,

$$\alpha^2\|y\|^2 + 2\alpha(x,y) + \|x\|^2 \geqslant 0$$

for every $\alpha \in \mathbb{R}$. Thus,

$$4|(x,y)|^2 \leqslant 4\|x\|^2\|y\|^2 \implies |(x,y)| \leqslant \|x\|\|y\|. \tag{3}$$

Finally, note that

$$\|x+y\|^2 = \|x\|^2 + \|y\|^2 + 2\Re(x,y) \leqslant \|x\|^2 + \|y\|^2 + 2|(x,y)| \leqslant (\|x\| + \|y\|)^2,$$

thereby completing the proof. ∎

THEOREM 7.3. Let H be a Hilbert space. Every nonempty closed convex $E \subseteq H$ contains a unique x of minimal norm.

Proof. Let

$$d = \inf\{\|x\| : x \in E\}.$$

Choose a sequence (x_n) in E such that $\|x_n\| \to d$ as $n \to \infty$. Since E is convex, $\frac{1}{2}(x_n + x_m) \in E$, whence $\|x_n + x_m\| \geqslant 2d$, for all $m, n \geqslant 1$.

Next, using the "parallelogram identity",

$$\|x_n - x_m\|^2 = 2\|x_n\|^2 + 2\|x_m\|^2 - \|x_n + x_m\|^2.$$

Let $\varepsilon > 0$ and choose $N \geqslant 1$ such that whenever $n \geqslant N$,

$$d \leqslant \|x_n\| \leqslant \sqrt{d^2 + \varepsilon^2}.$$

Thus, for $m, n \geqslant N$,

$$\|x_n - x_m\|^2 \leqslant 4d^2 + 4\varepsilon^2 - \|x_n + x_n\|^2 \leqslant 4\varepsilon^2,$$

thus $\|x_n - x_m\| \leqslant 2\varepsilon$ whenever $m, n \geqslant N$. This shows that (x_n) is Cauchy and hence, converges to some $x \in E$. Obviously, $\|x\| = d$.

As for uniqueness, suppose $x, y \in E$ with $\|x\| = \|y\| = d$. Then,

$$0 \leqslant \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 - \|x+y\|^2 \leqslant 2d^2 + 2d^2 - 4d^2 = 0.$$

Thus, $x = y$, thereby completing the proof. ∎

The above theorem fails quite spectacularly for Banach spaces.

EXAMPLE 7.4. Let $X = C[0,1]$ the $\mathbb{R}$-vector space of real-valued continuous functions on $[0,1]$ with the supremum norm. Let

$$M = \left\{ f \in X\colon \int_0^{1/2} f(t)\, dt - \int_{1/2}^1 f(t)\, dt = 1 \right\}.$$

Then, M is a closed convex subset of X but no element of M has minimal norm.

Proof. Obviously, M is convex. To see that it is closed, note that the linear functional

$$T : X \to \mathbb{R} \qquad f \mapsto \int_0^{1/2} f(t)\, dt - \int_{1/2}^1 f(t)\, dt$$

is a bounded linear functional, and hence, is continuous. Thus, M is closed too.

Next, for any $f \in M$,

$$1 = \left| \int_0^{1/2} f(t)\, dt - \int_{1/2}^1 f(t)\, dt \right| \leqslant \int_0^1 |f(t)|\, dt \leqslant \|f\|_\infty.$$

We contend that

$$\inf\{\|f\|_\infty \colon f \in M\} = 1.$$

To see this, fix some $0 < \delta < 1/2$. Define the function

$$f(x) = \begin{cases} 1+\varepsilon & 0 \leqslant x \leqslant \frac{1}{2} - \delta \\ \frac{1+\varepsilon}{\delta}\left(\frac{1}{2} - x\right) & \frac{1}{2} - \delta \leqslant x \leqslant \frac{1}{2} + \delta \\ -(1+\varepsilon) & \frac{1}{2} + \delta \leqslant x \leqslant 1. \end{cases}$$

Then,

$$\int_0^{1/2} f(t)\, dt - \int_{1/2}^1 f(t)\, dt = (1+\varepsilon)(1-2\delta) + \delta(1+\varepsilon) = (1-\delta)(1+\varepsilon).$$

Choosing

$$\varepsilon = \frac{\delta}{1-\delta},$$

we get $Tf = 1$. Note that $\|f\|_\infty = 1+\varepsilon$ and as $\delta \to 0^+$, we get $\|f\|_\infty \to 1^+$. This proves our claim.

Finally, suppose $f \in M$ such that $\|f\|_\infty = 1$. Then,

$$0 = \int_0^{1/2} 1 - f(t)\, dt + \int_{1/2}^1 1 + f(t)\, dt.$$

Since both integrals are non-negative and the functions are continuous, we must have $f(t) = 1$ whenever $0 \leqslant t \leqslant 1/2$ and $f(t) = -1$ whenever $1/2 \leqslant t \leqslant 1$, a contradiction. This completes the proof. ■

THEOREM 7.5. Let M be a closed subspace of a Hilbert space H, then $H = M \oplus M^{\perp}$.

Proof. Since

$$M^{\perp} = \bigcap_{x \in M} \ker(\cdot, x),$$

it is a closed subspace of H. Obviously, $M \cap M^{\perp} = \{0\}$. It remains to show that $H = M + M^{\perp}$. Indeed, let $x \in H$ and let $x_1 \in M$ be the unique element minimizing the distance to x. We contend that $x_2 = x - x_1$ is perpendicular to x_1.

Indeed, note that for every $y \in M$, we have

$$\|x_2\|^2 \leqslant \|x_2 + y\|^2 \implies \|y\|^2 + 2\Re(x_2, y) \geqslant 0,$$

for all $y \in M$. Suppose $(x_2, y) \neq 0$ for some $y \in M$. We can choose y such that $\Re(x_2, y) = -|(x_2, y)|$. Then, replacing y by αy for some $\alpha > 0$, we have $\alpha^2 \|y\|^2 - 2\alpha|(x, y)| \geqslant 0$ for all $\alpha > 0$. This is obviously false, and hence, $(x_2, y) \neq 0$ for all $y \in M$, thereby completing the proof. ∎

The above theorem fails for closed subspaces of Banach spaces.

EXAMPLE 7.6. $c_0 \subseteq \ell^\infty$ is not complemented.

Proof. We begin with a claim.
Claim. Let $T : \ell^\infty \to \ell^\infty$ be a bounded linear operator with $c_0 \subseteq \ker T$. Then there is an infinite subset $A \subseteq \mathbb{N}$ such that $Tx = 0$ whenever x is supported in A.
Proof of Claim: Suppose not. Then, for every infinite subset $A \subseteq \mathbb{N}$, there is an $x \in \ell^\infty$, supported in A such that $Tx \neq 0$. Choose an uncountable collection $\{A_i \colon i \in I\}$ of infinite subsets of $\mathbb{N}$ with pairwise finite intersections. According to our assumption, there are $x_i \in \ell^\infty$ supported in A_i with $Tx_i \neq 0$ and $\|x_i\| = 1$.

Since I is uncountable, there is an $n \in \mathbb{N}$ such that

$$I_n = \{i \in I \colon (Tx_i)(n) \neq 0\}$$

is uncountable (because the union of all the I_n's is I). Further, there is a positive integer k such that

$$I_{n,k} = \left\{ i \in I \colon |(Tx_i)(n)| \geqslant \frac{1}{k} \right\}$$

is uncountable (because the union of all the $I_{n,k}$'s is I_n).

Let $J \subseteq I_{n,k}$ be finite and set

$$y = \sum_{j \in J} \operatorname{sgn}\big((Tx_j)(n)\big) \cdot x_j.$$

Then,

$$(Ty)(n) = \sum_{j \in J} \operatorname{sgn}\big((Tx_j)(n)\big) \cdot (Tx_j)(n) \geqslant \sum_{j \in J} \frac{1}{k} = \frac{|J|}{k}.$$

Note that for $i \neq j$, $A_i \cap A_j$ is finite and hence, we can write $y = x + z$, where x has finite support and $\|z\| \leqslant 1$. Thus, $x \in c_0 \subseteq \ker T$ and hence,

$$\frac{|J|}{k} \leqslant \|Ty\| = \|Tx + Tz\| = \|Tz\| \leqslant \|T\| \implies |J| \leqslant k\|T\|,$$

which is absurd, since $I_{n,k}$ is infinite. This proves the claim. $\square$

Coming back, suppose c_0 were complemented in ℓ^∞. Then, there would be a projection operator $P : \ell^\infty \to c_0 \subseteq \ell^\infty$. Set $T = \mathbf{id} - P$. Since $c_0 \subseteq \ker T$, due to the claim above, there is an infinite subset $A \subseteq \mathbb{N}$, such that $Tx = 0$ whenever x is supported in A. Consider $\chi_A \in \ell^\infty$, the characteristic function of the set A. But note that

$$P(\chi_A) = (\mathbf{id} - T)(\chi_A) = \chi_A \notin c_0,$$

a contradiction. This completes the proof. ■

THEOREM 7.7 (RIESZ REPRESENTATION LEMMA). Let H be a Hilbert space. The natural map $H \to H^*$ given by $y \mapsto (\cdot, y)$ is an isometric and surjective.

Proof. Obviously, the map is injective and linear. To see isometry, note that $(y, y) = \|y\|^2$, whence $\|(\cdot, , y)\| \geqslant \|y\|$ and due to Cauchy-Schwarz,

$$|(x, y)| \leqslant \|x\|\|y\| \implies \|(\cdot, y)\| \leqslant \|y\| \implies \|(\cdot, y)\| = \|y\|.$$

It remains to show surjectivity. Let $\Lambda \neq 0$ be a continuous linear functional on H and $N = \ker \Lambda$. Since N is closed, we can write $H = N \oplus N^\perp$. Choose a nonzero vector $z \in N^\perp$. For any $x \in H$,

$$x - \frac{\Lambda x}{\Lambda z} z \in \ker \Lambda,$$

whence

$$0 = \left(x - \frac{\Lambda x}{\Lambda z} z, z\right) = (x, z) - \frac{\Lambda x}{\Lambda z}\|z\|^2.$$

Thus,

$$\Lambda x = \left(x, \frac{\overline{\Lambda z}}{\|z\|^2} z\right),$$

thereby completing the proof. ■

THEOREM 7.8. Let H be a Hilbert space and suppose $f : H \times H \to \mathbb{K}$ is sesquilinear and bounded, that is,

$$M := \sup\{|f(x, y)| \colon \|x\| = \|y\| = 1\} < \infty,$$

then there exists a unique $S \in \mathscr{B}(H)$ such that

$$f(x, y) = (x, Sy) \quad \forall x, y \in H.$$

Further, $\|S\| = M$.

Proof. Fix $y \in H$ and consider the mapping $x \mapsto f(x,y)$. This is a continuous linear functional on H and hence, is given by $x \mapsto (x, Sy)$ for a unique $Sy \in H$. We claim that the association $y \mapsto Sy$ is linear.

Indeed, if $y_1, y_2 \in H$, then

$$f(\cdot, y_1 + y_2) = f(\cdot, y_1) + f(\cdot, y_2) = f(\cdot, Sy_1) + f(\cdot, Sy_2) = f(\cdot, Sy_1 + Sy_2).$$

Due to uniqueness of $S(y_1 + y_2)$, we see that $S(y_1 + y_2) = Sy_1 + Sy_2$. Next, let $\alpha \in \mathbb{K}$ and $y \in H$. Then,

$$(\cdot, S(\alpha y)) = f(\cdot, \alpha y) = \overline{\alpha} f(\cdot, y) = \overline{\alpha}(\cdot, Sy) = (\cdot, \alpha Sy),$$

whence $S(\alpha y) = \alpha Sy$, i.e., S is linear.

Finally, we must show that $\|S\| = M$. Indeed, for $\|x\| = \|y\| = 1$:

$$|f(x,y)| \leqslant |(x, Sy)| \leqslant \|x\| \|Sy\| \leqslant \|S\|,$$

whence $M \leqslant \|S\|$. On the other hand, if $Sy \neq 0$, then

$$\|Sy\| = \left(\frac{Sy}{\|Sy\|}, Sy \right) = f\left(\frac{Sy}{\|Sy\|}, y \right) \leqslant M$$

Taking supremum over $\|y\| = 1$, we have that $\|S\| \leqslant M \leqslant \|S\|$, thereby completing the proof. ■

§§ Adjoints

Definition 7.9. Let $T \in \mathscr{B}(H)$. The map $f : H \times H \to \mathbb{K}$ given by $f(x,y) = (Tx, y)$, is a bounded sesquilinear form on H, whence, there is a $T^* \in \mathscr{B}(H)$ such that

$$(Tx, y) = f(x,y) = (x, T^*y) \quad \forall x, y \in H.$$

Next, note that

$$(x, Ty) = \overline{(y, T^*x)} = (T^*x, y) = (x, T^{**}y) \quad \forall x, y \in H.$$

Hence, $T^{**} = T$. On the other hand,

$$\|T^*\| = \sup\{|(Tx, y)| : \|x\| = \|y\| = 1\} \leqslant \|T\|.$$

Consequently, $\|T\| = \|T^{**}\| \leqslant \|T^*\| \leqslant \|T\|$, whence, $\|T^*\| = \|T\|$.

Similarly, the following identities are easy to show for $S, T \in \mathscr{B}(H)$:

$$(S+T)^* = S^* + T^*, \quad (\alpha S)^* = \overline{\alpha} S^*, \quad \text{and} \quad (ST)^* = T^*S^*.$$

Therefore,

$$\|Tx\|^2 = (Tx, Tx) = (x, T^*Tx) \leqslant \|T^*T\| \|x\|^2 \quad \forall x \in H.$$

Hence, $\|T\|^2 \leqslant \|T^*T\| \leqslant \|T^*\| \|T\| = \|T\|^2$, whence $\|T\|^2 = \|T^*T\|$. This makes $\mathscr{B}(H)$ a C*-algebra.

§§ Compact Self-Adjoint Operators

LEMMA 7.10. Let H be a Hilbert space and $T \in \mathscr{B}(H)$ a compact self-adjoint operator. Then

$$\|T\| = \sup\{|\langle Tx, x\rangle| : \|x\| = 1\}.$$

Proof. Let B denote the quantity on the right hand side. Due to the Cauchy-Schwarz Inequality, $B \leqslant \|T\|$. Let $x \neq 0$ and set $\lambda = \sqrt{\frac{\|Tx\|}{\|x\|}}$.

We have

$$\begin{aligned}
\langle Tx, Tx\rangle &= \frac{1}{4}\left|\langle T(\lambda x + \lambda^{-1}Tx), \lambda x + \lambda^{-1}Tx\rangle - \langle T(\lambda x - \lambda^{-1}Tx), \lambda x - \lambda^{-1}Tx\rangle\right| \\
&\leqslant \frac{1}{4}\left|\langle T(\lambda x + \lambda^{-1}Tx), \lambda x + \lambda^{-1}Tx\rangle\right| + \frac{1}{4}\left|\langle T(\lambda x + \lambda^{-1}Tx), \lambda x + \lambda^{-1}Tx\rangle\right| \\
&\leqslant \frac{B}{4}\left(\|\lambda x + \lambda^{-1}Tx\|^2 + \|\lambda x - \lambda^{-1}Tx\|^2\right) \\
&= \frac{B}{2}\left(\|\lambda x\|^2 + \|\lambda^{-1}Tx\|^2\right) \\
&= B\|x\|\|Tx\|.
\end{aligned}$$

Thus, $\|Tx\| \leqslant B\|x\|$, whence $\|T\| \leqslant B$, thereby completing the proof. ■

LEMMA 7.11. With the notation of the preceding lemma, either $\|T\|$ or $-\|T\|$ is an eigenvalue of T.

Proof. Due to the preceding lemma, there is a sequence of unit vectors (x_n) in H such that $|\langle Tx_n, x_n\rangle| \to \|T\|$. Since T is self-adjoint,

$$\overline{\langle Tx, x\rangle} = \langle x, Tx\rangle = \langle Tx, x\rangle,$$

and hence, $\langle Tx, x\rangle \in \mathbb{R}$. Therefore, moving to a subsequence, we may suppose that $\langle Tx_n, x_n\rangle \to \lambda \in \{\pm\|T\|\}$. Further, since T is compact, we may replace (x_n) with a subsequence such that $Tx_n \to \lambda y$ for some $y \in H$.

We contend that $x_n \to y$. First, note that

$$|\langle Tx_n, x_n\rangle| \leqslant \|Tx_n\|\|x_n\| = \|Tx_n\| \leqslant \|T\| = |\lambda|.$$

By our choice of the sequence (x_n), $|\langle Tx_n, x_n\rangle| \to |\lambda|$ and hence, $\|Tx_n\| \to |\lambda|$. Next,

$$\begin{aligned}
\|\lambda x_n - Tx_n\|^2 &= \langle \lambda x_n - Tx_n, \lambda x_n - Tx_n\rangle \\
&= \lambda^2 + \|Tx_n\|^2 - \langle \lambda x_n, Tx_n\rangle - \langle Tx_n, \lambda x_n\rangle \\
&= \lambda^2 + \|Tx_n\|^2 - 2\lambda\langle Tx_n, x_n\rangle
\end{aligned}$$

which goes to 0 as $n \to \infty$. Hence, $\|\lambda x_n - Tx_n\| \to 0$ as $n \to \infty$, consequently, $x_n \to y$, thereby completing the proof. ■

§8 Banach Algebras

Definition 8.1. A *Banach algebra* is a $\mathbb{C}$-algebra $\mathcal{A}$ equipped with a norm $\|\cdot\| : \mathcal{A} \to [0, \infty)$ with respect to which it is a Banach space and

$$\|xy\| \leqslant \|x\|\|y\| \quad \forall x, y \in \mathcal{A}.$$

The Banach algebra is said to be *unital* if it possesses a multiplicative identity.

An *involution* on an algebra $\mathcal{A}$ is a map

$$\mathcal{A} \to \mathcal{A} \quad x \mapsto x^*$$

of order 2 that satisfies

$$(x+y)^* = x^* + y^* \quad (\lambda x)^* = \overline{\lambda} x^* \quad (xy)^* = y^* x^*.$$

An algebra equipped with such an involution is called a *$*$-algebra*. A Banach $*$-algebra that satisfies

$$\|x^* x\| = \|x\|^2 \quad \forall x \in \mathcal{A}$$

is called a *C^*-algebra*.

Remark 8.2. If $\mathcal{A}$ is a C^*-algebra, for $x \neq 0$, we have

$$\|x\|^2 = \|x^* x\| \leqslant \|x^*\|\|x\| \implies \|x\| \leqslant \|x^*\| \leqslant \|x^{**}\| = \|x\|,$$

whence $\|x\| = \|x^*\|$. That is, the involution is an isometry.

Definition 8.3. If $\mathcal{A}$ and $\mathcal{B}$ are Banach algebras, a *homomorphism* is a bounded linear map $\phi : \mathcal{A} \to \mathcal{B}$ such that $\phi(xy) = \phi(x)\phi(y)$ for all $x, y \in \mathcal{A}$.

Further, if $\mathcal{A}$ and $\mathcal{B}$ are Banach $*$-algebras, a *$*$-homomorphism* is a homomorphism of Banach algebras $\phi : \mathcal{A} \to \mathcal{B}$ such that $\phi(x^*) = \phi(x)^*$ for all $x \in \mathcal{A}$.

Theorem 8.4. Let $\mathcal{A}$ be a unital Banach algebra.

(a) If $|\lambda| > \|x\|$, then $\lambda - x$ is invertible in $\mathcal{A}$.

(b) If x is invertible, and $\|y\| < \|x^{-1}\|^{-1}$, then $x - y$ is invertible with inverse

$$(x-y)^{-1} = \sum_{n \geqslant 0} (x^{-1} y)^n x^{-1}.$$

(c) If x is invertible and $\|y\| < \frac{1}{2}\|x^{-1}\|^{-1}$, then

$$\|(x-y)^{-1} - x^{-1}\| < 2\|x^{-1}\|^2 \|y\|.$$

(d) $\mathcal{A}^\times \subseteq \mathcal{A}$ is open and $x \mapsto x^{-1}$ on $\mathcal{A}^\times$ is continuous.

Proof. (a) We have

$$(\lambda - x)^{-1} = \lambda^{-1}\left(e - \lambda^{-1}x\right)^{-1} = \sum_{n \geqslant 0} \lambda^{-(n+1)} x^{-n},$$

which converges because things are Cauchy and all the good stuff.

(b) Again, we can write

$$(x-y)^{-1} = \left(x(e - x^{-1}y)\right)^{-1} = (e - x^{-1}y)^{-1}x^{-1} = \sum_{n \geqslant 0} (x^{-1}y)x^{-1}.$$

(c) Using the above expansion, we can write

$$\|(x-y)^{-1} - x^{-1}\| \leqslant \sum_{n \geqslant 0} \|x^{-1}\|^{n+2}\|y\|^{n+1} < 2\|x^{-1}\|^2\|y\|.$$

(d) Due to part (b), $\mathcal{A}^\times$ is open in $\mathcal{A}$ and due to part (c), $x \mapsto x^{-1}$ is continuous. ∎

DEFINITION 8.5. Let $\mathcal{A}$ be a unital Banach algebra and $x \in \mathcal{A}$. The *spectrum* of x is

$$\sigma(x) = \{\lambda \in \mathbb{C} \colon \lambda e - x \text{ is not invertible}\}.$$

For $\lambda \notin \sigma(x)$, define the *resolvent* of x as

$$R_x(\lambda) = (\lambda e - x)^{-1} : \mathbb{C} \setminus \sigma(x) \to \mathcal{A}.$$

PROPOSITION 8.6. For any $x \in \mathcal{A}$, $\sigma(x)$ is a compact subset of $\mathbb{C}$ that is contained in the disk $\{\lambda \in \mathbb{C} \colon |\lambda| \leqslant \|x\|\}$.

Proof. Obviously, if $|\lambda| > \|x\|$, then $\lambda e - x$ is invertible. Thus, $\sigma(x)$ is contained in the above disk. Consider the map $\lambda \mapsto \lambda e - x$, which is continuous and hence, the preimage of $\mathcal{A}^\times$ is open in $\mathbb{C}$. As a result, $\sigma(x)$ is closed, thereby completing the proof. ∎

PROPOSITION 8.7. R_x is an analytic function. And, $R_x(\lambda) \to 0$ as $\lambda \to \infty$.

Proof. We have

$$\begin{aligned} R_x(\mu) - R_x(\lambda) &= (\mu e - x)^{-1} - (\lambda e - x)^{-1} \\ &= R_x(\mu)\left((\lambda e - x) - (\mu e - x)\right) R_x(\lambda). \end{aligned}$$

Hence,

$$\frac{R_x(\mu) - R_x(\lambda)}{\mu - \lambda} = -R_x(\mu)R_x(\lambda).$$

In the limit $\mu \to \lambda$, we get

$$R_x'(\lambda) = -R_x(\lambda)^2.$$

As for the second part, simply note that for $|\lambda| > \|x\|$,

$$\|R_x(\lambda)\| = \left\|\sum_{n \geqslant 0} \lambda^{-(n+1)} x^n\right\| \leqslant |\lambda|^{-1} \sum_{n \geqslant 0} |\lambda|^{-n}\|x\|^n = \frac{1}{|\lambda| - \|x\|},$$

which goes to 0 as $\lambda \to \infty$, thereby completing the proof. ∎

THEOREM 8.8 (GELFAND-MAZUR). Let $\mathcal{A}$ be a unital Banach algebra $\sigma(x) \neq \emptyset$ for every $x \in \mathcal{A}$.

Proof. Suppose $\sigma(x) = \emptyset$ for some $x \in \mathcal{A}$. Then, $R_x : \mathbb{C} \to \mathcal{A}$ is an analytic function. For any $\Lambda \in \mathcal{A}^*$, $\Lambda \circ R_x$ is an entire function and is bounded, since

$$\lim_{\lambda \to \infty} \Lambda(R_x(\lambda)) = \Lambda\left(\lim_{\lambda \to \infty} R_x(\lambda)\right) = 0.$$

Due to Liouville's Theorem, $\Lambda \circ R_x$ must be constant on $\mathbb{C}$ and equal to 0. Since this is true for every $\Lambda \in \mathcal{A}^*$, we see that $R(\lambda) = 0$ for every $\lambda \in \mathbb{C}$, which is absurd. This completes the proof. ∎

COROLLARY. If $\mathcal{A}$ is a unital Banach algebra in which every nonzero element is invertible, then $\mathcal{A} = \mathbb{C}e$.

Proof. Suppose $x \in \mathcal{A} \setminus \mathbb{C}e$, then $\lambda e - x \neq 0$ for every $\lambda \in \mathbb{C}$, whence, $\lambda e - x$ is invertible for every $\lambda \in \mathbb{C}$, a contradiction. ∎

DEFINITION 8.9. Let $\mathcal{A}$ be a unital Banach algebra. For $x \in \mathcal{A}$, the *spectral radius* of x is defined to be

$$\rho(x) := \sup\{|\lambda| \colon \lambda \in \sigma(x)\}.$$

We have the obvious inequality $\rho(x) \leqslant \|x\|$.

THEOREM 8.10 (SPECTRAL RADIUS FORMULA). Let $\mathcal{A}$ be a unital Banach algebra and $x \in \mathcal{A}$. Then,

$$\rho(x) = \lim_{n \to \infty} \|x^n\|^{1/n}.$$

Proof. If $\lambda \in \sigma(x)$, then

$$\lambda^n e - x^n = (\lambda e - x)\left(\lambda^{n-1} e + \cdots + x^{n-1}\right).$$

Consequently, $\lambda^n e - x^n$ cannot be invertible. Hence, $|\lambda|^n \leqslant \|x^n\|$. In particular, this gives

$$\rho(x) \leqslant \liminf_{n \to \infty} \|x^n\|^{1/n}.$$

Next, for $|\lambda| > \|x\|$, we have a Laurent series about infinity:

$$\Lambda \circ R_x(\lambda) = \sum_{n \geqslant 0} \lambda^{-(n+1)} \Lambda(x^n).$$

Note that $\Lambda \circ R_x$ is analytic on $|\lambda| > \rho(x)$ and hence, the above Laurent series must be valid there too.

Hence, for any $|\lambda| > \rho(x)$, there is a constant $C_\Lambda > 0$ such that

$$|\Lambda(\lambda^{-n} x^n)| = |\lambda^{-n} \Lambda(x^n)| \leqslant C_\Lambda \quad \forall n \in \mathbb{N}.$$

This holds for all $\Lambda \in \mathcal{A}^*$. Thus, the sequence $(\lambda^{-n}x^n)$ is bounded, that is, there is a $C > 0$ such that $\|x^n\| \leqslant C|\lambda|^n$. Hence,

$$\limsup_{n\to\infty} \|x^n\|^{1/n} \leqslant \limsup_{n\to\infty} C^{1/n}|\lambda| = |\lambda|.$$

Taking infimum over λ, we get

$$\limsup_{n\to\infty} \|x^n\|^{1/n} \leqslant \rho(x) \leqslant \liminf_{n\to\infty} \|x^n\|^{1/n},$$

thereby completing the proof. ∎

DEFINITION 8.11. Let $\mathcal{A}$ be a unital Banach algebra. A *multiplicative functional* on $\mathcal{A}$ is a *nonzero* homomorhpism $h : \mathcal{A} \to \mathbb{C}$. The set of all multiplicative functionals on $\mathcal{A}$ is called the *spectrum* of $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$.

PROPOSITION 8.12. Let $\mathcal{A}$ be a unital Banach algebra and suppose $h \in \sigma(\mathcal{A})$.

(a) $h(e) = 1$.

(b) If $x \in \mathcal{A}^\times$, then $h(x) \neq 0$.

(c) $|h(x)| \leqslant \rho(x) \leqslant \|x\|$ for all $x \in \mathcal{A}$. That is, $\|h\| \leqslant 1$.

Proof. (a) Since $h \neq 0$, there is an $x \in \mathcal{A}$ such that $h(x) \neq 0$. Then,

$$h(x) = h(xe) = h(x)h(e) \implies h(e) = 1.$$

(b) Obviously,

$$1 = h(e) = h(x^{-1}x) = h(x^{-1})h(x) \implies h(x) \neq 0.$$

(c) Suppose $|\lambda| > \rho(x)$. Then, $\lambda e - x \in \mathcal{A}^\times$, consequently,

$$0 \neq h(\lambda e - x) = \lambda - h(x) \implies h(x) \neq \lambda.$$

Since this holds for all $|\lambda| > \rho(x)$, we have $|h(x)| \leqslant \rho(x) \leqslant \|x\|$. ∎

As a consequence, $\sigma(\mathcal{A})$ is contained in the closed unit ball of $\mathcal{A}^*$. Equip the latter with the weak*-topology. Using nets, it is easy to see that $\sigma(\mathcal{A})$ is closed in $\mathcal{A}^*$. Due to Banach-Alaoglu, the closed unit ball of $\mathcal{A}^*$ is weak*-compact and hence, so is $\sigma(\mathcal{A})$ with the subspace topology from the weak*-topology on $\mathcal{A}^*$. Thus, $\sigma(\mathcal{A})$ is a *compact Hausdorff space*.

PROPOSITION 8.13. Let $\mathcal{A}$ be a commutative unital Banach algebra and $\mathcal{J} \subsetneq \mathcal{A}$ be a proper ideal.

(a) $\mathcal{J} \subseteq \mathcal{A} \setminus \mathcal{A}^\times$

(b) $\overline{\mathcal{J}}$ is a proper ideal.

(c) $\mathcal{J}$ is contained in a maximal ideal.

(d) Every maximal ideal is closed.

Proof. The first assertion is obvious. As for the second, note that $\mathcal{A} \setminus \mathcal{A}^{\times}$ is closed and hence, $\overline{\mathcal{J}} \subseteq \mathcal{A} \setminus \mathcal{A}^{\times}$. Consequently, $\overline{\mathcal{J}} \neq \mathcal{A}$. To see that it is an ideal, suppose $x \in \overline{\mathcal{J}}$ and $a \in \mathcal{A}$. Then, there is a sequence (x_n) converging to x. Consequently, (ax_n) converges to ax. But each $ax_n \in \mathcal{J}$ and hence, $ax \in \overline{\mathcal{J}}$. This proves (b).

The third assertion is a standard application of Zorn's lemma. As for (d), if $\mathcal{M}$ is a maximal ideal, then $\mathcal{M} \subseteq \overline{\mathcal{M}} \subsetneq \mathcal{A}$ due to (b). The maximality of $\mathcal{M}$ forces $\mathcal{M} = \overline{\mathcal{M}}$, thereby completing the proof. ∎

THEOREM 8.14. Let $\mathcal{A}$ be a commutative unital Banach algebra. Then, the map $h \mapsto \ker h$ is a bijective correspondence between $\sigma(\mathcal{A})$ and the set of all maximal ideals in $\mathcal{A}$.

Proof. The map is obviously an injection. We establish surjectivity. Let $\mathcal{M}$ be a maximal ideal in $\mathcal{A}$ and consider the quotient algebra $\mathcal{A}/\mathcal{M}$ equipped with the norm:

$$\|x + \mathcal{M}\| = \inf\{\|x + y\| : y \in \mathcal{M}\}.$$

This is again a commutative unital Banach algebra in which every non-zero element is invertible (standard fact from ring theory). Due to Gelfand-Mazur, $\mathcal{A}/\mathcal{M} \cong \mathbb{C}(e + \mathcal{M})$. The composition

$$\mathcal{A} \longrightarrow \mathcal{A}/\mathcal{M} \cong \mathbb{C}(e + \mathcal{M}) \cong \mathbb{C}$$

is the required linear functional, thereby proving surjectivity. ∎

DEFINITION 8.15. Let $\mathcal{A}$ be a commutative unital Banach algebra. For each $x \in \mathcal{A}$, there is a continuous function $\widehat{x} : \sigma(\mathcal{A}) \to \mathbb{C}$ given by $h \mapsto h(x)$. This gives a map

$$\Gamma_{\mathcal{A}} : \mathcal{A} \to C(\sigma(A)) \qquad x \mapsto \widehat{x},$$

known as the *Gelfand transform* on $\mathcal{A}$.

PROPOSITION 8.16. Let $\mathcal{A}$ be a commutative unital Banach algebra and $x \in \mathcal{A}$.

(a) The $\Gamma : \mathcal{A} \to C(\sigma(\mathcal{A}))$ is a homomorphism, and $\widehat{e}$ is the constant function 1.

(b) x is invertible if and only if $\widehat{x}$ never vanishes.

(c) The range of $\widehat{x} : \sigma(\mathcal{A}) \to \mathbb{C}$ is precisely $\sigma(x)$.

(d) $\|\widehat{x}\|_{\sup} = \rho(x) \leqslant \|x\|$.

Proof. (a) Obvious.

(b) If x is invertible, then due to (a), so is $\widehat{x}$, whence it never vanishes. On the other hand, if x is not invertible, then it is contained in some maximal ideal $\mathfrak{M}$, whence, there is an $h \in \sigma(\mathcal{A})$ that vanishes on x. Thus, $\widehat{x}(h) = 0$, that is, $\widehat{x}$ vanishes somewhere.

(c) Next, suppose $\lambda = \widehat{x}(h) = h(x)$. Then, $h(\lambda e - x) = 0$, hence, $\lambda e - x$ is not invertible, i.e. $\lambda \in \sigma(x)$. Similarly, if $\lambda \in \sigma(x)$, then $\lambda e - x$ is not invertible and hence, $\widehat{x}$ vanishes somewhere, consequently, $h(\lambda e - x) = 0$ for some $h \in \sigma(\mathcal{A})$. This shows that λ is in the range of $\widehat{x}$.

(d) Follows from (c). ■

DEFINITION 8.17. Let $\mathcal{A}$ be a commutative unital Banach $*$-algebra. If $\Gamma : \mathcal{A} \to C(\sigma(\mathcal{A}))$ is a $*$-homomorphism, then $\mathcal{A}$ is said to be *symmetric*.

REMARK 8.18. Note that $\mathcal{A}$ being symmetric is the same as saying

$$\widehat{x^*} = \overline{\widehat{x}} \quad \forall x \in \mathcal{A}.$$

PROPOSITION 8.19. Let $\mathcal{A}$ be a commutative Banach $*$-algebra.

(a) $\mathcal{A}$ is symmetric if and only if $\widehat{x}$ is real-valued whenever $x = x^*$.

(b) Every C*-algebra is symmetric.

(c) If $\mathcal{A}$ is symmetric, $\Gamma(\mathcal{A})$ is dense in $C(\sigma(\mathcal{A}))$.

Proof. (a) If $\mathcal{A}$ is symmetric and $x^* = x$, then $\widehat{x} = \widehat{x^*} = \overline{\widehat{x}}$, whence $\widehat{x}$ is real-valued. Next, we prove the converse. For any $x \in \mathcal{A}$, write

$$x = \underbrace{\frac{x + x^*}{2}}_{y} + \underbrace{\frac{x - x^*}{2}}_{z}.$$

Note that $y^* = y$ and $z + z^* = 0$. Our hypothesis implies $\widehat{y}$ is real-valued and $\widehat{z} + \overline{\widehat{z}} = 0$. Thus,

$$\widehat{x^*} = \widehat{y^*} + \widehat{z^*} = \widehat{y} - \widehat{z} = \widehat{y} + \overline{\widehat{z}} = \overline{\widehat{x}}.$$

(b) Let $x \in \mathcal{A}$ be such that $x^* = x$. Suppose $h(x) = \alpha + i\beta$. We shall show that $\beta = 0$. Indeed, for $t \in \mathbb{R}$, let $z = x + ite$. Then,

$$z^*z = (x - ite)(x + ite) = x^2 + t^2 e.$$

And hence,

$$|\alpha + (\beta + t)i|^2 = |h(z)|^2 \leqslant \|z\|^2 = \|z^*z\| = \|x^2 + t^2 e\| \leqslant \|x\|^2 + t^2.$$

That is,

$$\alpha^2 + 2\beta t + \beta^2 \leqslant \|x\|^2 \quad \forall t \in \mathbb{R}.$$

Thus, $\beta = 0$ and due to (a), $\mathcal{A}$ is symmetric.

(c) Note that $\Gamma(\mathcal{A})$ contains all the constant functions and thus, the family $\Gamma(\mathcal{A})$ does not vanish at any point. Next, by definition, $\Gamma(\mathcal{A})$ separates points. Further, since Γ is a $*$-homomorophism, $\Gamma(\mathcal{A})$ is closed under taking conjugates. Thus, $\Gamma(\mathcal{A})$ is dense in $C(\sigma(\mathcal{A}))$ due to the Stone-Weierstrass Theorem. ■

PROPOSITION 8.20. Let $\mathcal{A}$ be a commutative unital Banach algebra.

(a) If $x \in \mathcal{A}$, then $\|\widehat{x}\|_{\sup} = \|x\|$ if and only if $\|x^{2^k}\| = \|x\|^{2^k}$ for all $k \geqslant 1$.

(b) $\Gamma : \mathcal{A} \to C(\sigma(\mathcal{A}))$ is an isometry if and only if $\|x^2\| = \|x\|^2$ for all $x \in \mathcal{A}$.

Proof. (a) This follows immediately from the spectral radius formula.

$$\|\widehat{x}\|_{\sup} = \rho(x) = \lim_{k\to\infty} \|x^{2^k}\|^{1/2^k} = \lim_{k\to\infty} \|x\|^{2^k \cdot 2^{-k}} = \|x\|.$$

(b) We have

$$\|x^{2^k}\| = \|x^{2^{k-1}}\|^2 = \cdots = \|x\|^{2^k}.$$

■

THEOREM 8.21 (GELFAND-NAIMARK). If $\mathcal{A}$ is a commutative unital C*-algebra, then $\Gamma : \mathcal{A} \to C(\sigma(\mathcal{A}))$ is an isometric $*$-isomorphism.

Proof. That Γ is a $*$-homomorphism has already been established. We first show that Γ is an isometry. Let $x \in \mathcal{A}$ and set $y = x^* x$. Then, $y^* = y$, so

$$\|y^{2^k}\| = \left\| \left(y^{2^{k-1}}\right)^* y^{2^{k-1}} \right\| = \|y^{2^{k-1}}\|^2 = \cdots = \|y\|^{2^k}.$$

Due to part (a) of the preceding result, $\|\widehat{y}\|_{\sup} = \|y\|$. But $\widehat{y} = \overline{\widehat{x}}\widehat{x} = |\widehat{x}|^2$. Hence,

$$\|\widehat{x}\|_{\sup}^2 = \|\widehat{y}\|_{\sup} = \|y\| = \|x\|^2 \implies \|\widehat{x}\|_{\sup} = \|x\|,$$

whence, due to part (b) of the preceding result, Γ is an isometry. Thus, its image is closed in $C(\sigma(\mathcal{A}))$. But we already argued that $\Gamma(\mathcal{A})$ is dense in $C(\sigma(\mathcal{A}))$ and hence, Γ must be surjective. This completes the proof. ■

§9 DISTRIBUTIONS

§§ The topology on $\mathscr{D}_K$

Let $\Omega \subseteq \mathbb{R}^n$ be an open set. We begin by topologizing $C^\infty(\Omega)$. Fix an exhaustion $\{K_i\}$ of Ω by compact sets. That is,

- $\Omega = \bigcup_{i=1}^{\infty} K_i$, and
- $K_i \subseteq K_{i+1}^\circ$ for all $i \geqslant 1$.

Define the seminorms $p_N : C^\infty(\Omega) \to \mathbb{R}$ given by

$$p_N(\phi) = \sup \{ |\partial^\alpha \phi(x)| : x \in K_N,\ |\alpha| \leqslant N \}.$$

That this is a separating family of seminorms is obvious, and since this is a countable family, the induced locally convex vector topology on $C^\infty(\Omega)$ is metrizable.

It is easy to see that the "evaluation functionals" on $C^\infty(\Omega)$ equipped with this topology are continuous, therefore,

$$\mathscr{D}_K := \bigcap_{x\in\Omega\setminus K} \ker \mathrm{ev}_x$$

is closed in $C^\infty(\Omega)$. It is easy to see that a (countable) local base at 0 is given by the sets

$$V_N = \left\{ f \in C^\infty(\Omega) \colon p_N(f) < \frac{1}{N} \right\}$$

for $N \geqslant 1$. Further, in this topology $C^\infty(\Omega)$ is a Fréchet space[1] and since $\mathscr{D}_K$ is closed, it to is a Fréchet space. It can also be showed that $C^\infty(\Omega)$ has the Heine-Borel property and hence, the same conclusion holds for $\mathscr{D}_K$.

§§ Distributions

Let $\Omega \subseteq \mathbb{R}^n$ be an open set. Define

$$\mathscr{D}(\Omega) = \bigcup_{K\Subset\Omega} \mathscr{D}_K.$$

Introduce the seminorms $\|\cdot\|_N : \mathscr{D}(\Omega) \to \mathbb{R}$ given by

$$\|\phi\|_N = \max\left\{ |\partial^\alpha \phi(x)| \colon x \in \Omega,\ |\alpha| \leqslant N \right\},$$

for $\phi \in \mathscr{D}(\Omega)$ and $N \geqslant 0$. The restrictions of these seminorms to $\mathscr{D}_K$ are still seminorms. We claim that they induce the same topology as the canoical topology of $\mathscr{D}_K$ discussed in the preceding (sub)section. First, there is a positive integer N_0 such that $K \subseteq K_N$ for all $N \geqslant N_0$. For these N, $\|\phi\|_N = p_N(\phi)$ if $\phi \in \mathscr{D}_K$. Further, since $\|\phi\|_N \leqslant \|\phi\|_{N+1}$, the topologies induced by either sequence of seminorms are unchanged if we let N start at N_0 instead of 1. Thus, the two topologies coincide and a local base is given by sets of the form

$$V_N = \left\{ \phi \in \mathscr{D}_K \colon \|\phi\|_N < \frac{1}{N} \right\}.$$

In particular, $\mathscr{D}_K$ is still a Fréchet space having the Heine-Borel property.

DEFINITION 9.1. Let $\Omega \subseteq \mathbb{R}^n$ be a nonempty open set.

(a) For every compact $K \Subset \Omega$, let τ_K denote the standard Fréchet space topology of $\mathscr{D}_K$.

(b) Let β denote the collection of all convex balanced sets $W \subseteq \mathscr{D}(\Omega)$ such that $\mathscr{D}_K \cap W \in \tau_K$ for every $K \Subset \Omega$.

(c) τ is the collection of all unions of sets of the form $\phi + W$, with $\phi \in \mathscr{D}(\Omega)$ and $W \in \beta$.

THEOREM 9.2. (a) τ is a topology on $\mathscr{D}(\Omega)$, and β is a local base for τ.

[1] I might add in the details to this some day.

(b) τ makes $\mathscr{D}(\Omega)$ into a locally convex topological vector space.

Proof. (a) Let $V_1, V_2 \in \tau$. We shall show that for all $\phi \in V_1 \cap V_2$, there is some $W \in \beta$ such that $\phi + W \subseteq V_1 \cap V_2$. Since each V_i is open, there is some $W_i \in \beta$ such that $\phi \in \phi_i + W_i \subseteq$. Let $K \Subset \Omega$ such that $\phi, \phi_1, \phi_2 \in \mathscr{D}_K$. Since each $\mathscr{D}_K \cap W_i$ is open in $\mathscr{D}_K$, W_i is convex and balanced, and $\phi - \phi_i \in \mathscr{D}_K \cap W_i$. Since the Minkowski functional on $\mathscr{D}_K$ corresponding to $\mathscr{D}_K \cap W_i$ is continuous, there is a $0 < \delta_i < 1$ such that $\phi - \phi_i \in (1 - \delta_i)W_i$. Hence,

$$\phi - \phi_i + \delta_i W_i \subseteq (1 - \delta_i) \subseteq W_i \implies \phi + \delta_i W_i \subseteq \phi_i + W_i \subseteq V_i,$$

whence $\phi + (\delta_1 W_1) \cap (\delta_2 W_2) \subseteq V_1 \cap V_2$. Since $\delta_1 W_1 \cap \delta_2 W_2 \in \beta$, the conclusion follows.

(b) Since β consists of convex sets, it suffices to show that τ makes $\mathscr{D}(\Omega)$ a topological vector space. First, we must show that the space is T_1. Let $\phi_1 \neq \phi_2 \in \mathscr{D}(\Omega)$, and set

$$W = \{\phi \in \mathscr{D}(\Omega) \colon \|\phi\|_0 < \|\phi_1 - \phi_2\|_0\},$$

where $\|\cdot\|_0$ is precisely the sup-norm on Ω. By definition, it is easy to see that $W \in \beta$ and $\phi_1 \notin \phi_2 + W$, consequently, $\{\phi_1\}$ is closed.

To see that addition is continuous, let $(\phi_1, \phi_2) \mapsto \phi_1 + \phi_2$ and V an open set containing $\phi_1 + \phi_2$. Since β forms a local base for the topology, we can find some $W \in \beta$ such that $(\phi_1 + \phi_2) + W \subseteq V$, and

$$\left(\phi_1 + \frac{1}{2}W\right) + \left(\phi_2 + \frac{1}{2}W\right) \subseteq (\phi_1 + \phi_2) + W \subseteq V.$$

Thus, addition is continuous.

Finally, we must show that scalar multiplication is continuous. Let $\alpha_0 \in \mathbb{K}$ and $\phi_0 \in \mathscr{D}(\Omega)$. Then,

$$\alpha\phi - \alpha_0\phi_0 = \alpha(\phi - \phi_0) + (\alpha - \alpha_0)\phi_0.$$

Let V be an open set containing $\alpha_0\phi_0$, and choose a $W \in \beta$ such that $\alpha_0\phi_0 + W \subseteq V$. There is a $\delta > 0$ such that $\delta\phi_0 \in \frac{1}{2}W$. Next, choose $c > 0$ such that $2c(|\alpha_0| + \delta) = 1$. For $|\alpha - \alpha_0| < \delta$ and $\phi - \phi_0 \in cW$, we have

$$\alpha\phi - \alpha_0\phi_0 \in |\alpha|cW + \frac{1}{2}W \subseteq c(|\alpha_0| + \delta)W + \frac{1}{2}W \subseteq W,$$

as desired. This completes the proof. ■

THEOREM 9.3. (a) A convex balanced subset V of $\mathscr{D}(\Omega)$ is open if and only if $V \in \beta$.

(b) The topology τ_K of any $\mathscr{D}_K \subseteq \mathscr{D}(\Omega)$ coincides with the subspace topology that $\mathscr{D}_K$ inherits from $\mathscr{D}(\Omega)$.

(c) If E is a bounded subset of $\mathscr{D}(\Omega)$, then $E \subseteq \mathscr{D}_K$ for some $K \subseteq \Omega$ and there are real numbers $0 < M_N < \infty$ such that every $\phi \in E$ satisfies the inequalities $\|\phi\|_N \leqslant M_N$ for $N \geqslant 0$.

(d) $\mathscr{D}(\Omega)$ has the Heine-Borel property, that is, closed and bounded sets are compact.

(e) If $\{\phi_i\}$ is a Cauchy sequence in $\mathscr{D}(\Omega)$, then $\{\phi_i\} \subseteq \mathscr{D}_K$ for some compact $K \subseteq \Omega$, and

$$\lim_{(i,j)\to\infty} \|\phi_i - \phi_j\|_N = 0$$

for all $N \geqslant 0$.

(f) If $\phi_i \to 0$ in the topology of $\mathscr{D}(\Omega)$, then there is a compact set $K \subseteq \Omega$ which contains the support of every ϕ_i and $\partial^\alpha \phi_i \to 0$ uniformly as $i \to \infty$, for every multi-index α.

(g) $\mathscr{D}(\Omega)$ is a Fréchet space.

Proof. Let $V \in \tau$ and $\phi \in \mathscr{D}_K \cap V$. Since β form a local base, there is a $W \in \beta$ such that $\phi + W \subseteq V$. Hence,

$$\phi + (\mathscr{D}_K \cap W) \subseteq \mathscr{D}_K \cap V.$$

Since $\mathscr{D}_K \cap W$ is open in $\mathscr{D}_K$, we have that $\mathscr{D}_K \cap V \in \tau_K$.

(a) Now, let V be a convex balanced subset of $\mathscr{D}(\Omega)$. If V is open, then due to our observation above, $V \in \beta$. The converse direction is trivial since $\beta \subseteq \tau$.

(b) The above remark shows that the induced topology on $\mathscr{D}_K$ is coarser than τ_K. Conversely, suppose $E \in \tau_K$. We have to show that $E = \mathscr{D}_K \cap V$ for some $V|in\tau$. By definition, for every $\phi \in E$, there is a positive integer N and $\delta > 0$ such that

$$\{\psi \in \mathscr{D}_K \colon \|\psi - \phi\|_N < \delta\} \subseteq E.$$

Set $W_\phi = \{\psi \in \mathscr{D}(\Omega) \colon \|\psi\|_N < \delta\} \in \beta$, so that

$$\mathscr{D}_K \cap (\phi + W_\phi) = \phi + \mathscr{D}_K \cap W_\phi \subseteq E.$$

Since $W_\phi \in \beta$ for every $\phi \in E$, we see that $V := \bigcup\limits_{\phi \in E} (\phi + W_\phi)$ is an element of τ and $V \cap E = E$, as desired.

(c) Suppose E does not lies in any $\mathscr{D}_K$. Using an exhaustion of Ω, we can find a sequence of functions $\phi_m \in E$ and distinct points $x_m \in \Omega$ with no limit point in Ω such that $\phi_m(x_m) \neq 0$. Let W be the set of all $\phi \in \mathscr{D}(\Omega)$ which satisfy

$$|\phi(x_m)| < \frac{1}{m}|\phi_m(x_m)| \qquad \forall\, m \geqslant 1.$$

Note that

$$W \cap \mathscr{D}_K = \bigcap_{x_m \in W \cap \mathscr{D}_K} \left\{\phi \in \mathscr{D}_K \colon |\phi(x_m)| < \frac{1}{m}|\phi_m(x_m)|\right\},$$

which is a finite intersection since only finitely many of the x_m's can be contained in K (as they do not admit a limit point in Ω). Thus, $W \cap \mathscr{D}_K$ is open, owing to the continuity of the "evaluation functionals" on $\mathscr{D}_K$; hence $W \in \beta$. Since $\phi_m \notin mW$, no multiple of W contains E, which shows that E is not bounded. Hence, every bounded E lies in some $\mathscr{D}_K$. Being a bounded subset of $\mathscr{D}_K$, every seminorm on $\mathscr{D}_K$ is bounded on E, whence the last assertion of (c) follows.

(d) This follows immediately from the above parts, since every bounded set is contained in some $\mathscr{D}_K$, whose subspace topology is same as the canonical topology, in which it has the Heine-Borel property.

(e) Every Cauchy sequence is bounded and hence, is contained in some $\mathscr{D}_K$, which has its canonical topology induced by the seminorms $\|\cdot\|_N$, whence the conclusion follows.

(f) This follows immediately from (e).

(g) Finally, we have shown that any Cauchy sequence in $\mathscr{D}(\Omega)$ lies in $\mathscr{D}_K$, which is Fréchet, whence it must converge. This completes the proof. ■

THEOREM 9.4. Let Λ be a linear map from $\mathscr{D}(\Omega)$ to a locally convex space Y. Then the following are equivalent:

(a) Λ is continuous.

(b) Λ is bounded.

(c) If $\phi_i \to 0$ in $\mathscr{D}(\Omega)$, then $\Lambda\phi_i \to 0$ in Y.

(d) The restriction of Λ to every $\mathscr{D}_K \subseteq \mathscr{D}(\Omega)$ are continuous.

Proof. $(a) \implies (b)$ is well known. Next, if $\phi_i \to 0$ in $\mathscr{D}(\Omega)$, then it is contained in some $\mathscr{D}_K$ for a compact $K \Subset \Omega$. Since the restriction of Λ to $\mathscr{D}_K$ is continuous and it is a metrizable topological vector space, $\Lambda\phi_i \to 0$ in Y, thereby proving $(b) \implies (c)$.

To see $(c) \implies (d)$, it suffices to show that the restriction of Λ to each $\mathscr{D}_K$ is sequentially continuous. If $\phi_i \to 0$ in $\mathscr{D}_K$ and since the topology of $\mathscr{D}_K$ is the subspace topology, we see that $\phi_i \to 0$ in $\mathscr{D}(\Omega)$ and according to our assumption, $\Lambda\phi_i \to 0$ in Y, which proves sequential continuity.

Finally, let U be a convex balanced neighborhood of 0 in Y. It suffices to show that $V = \Lambda^{-1}U$ is open. Note that V is a convex balanced subset of $\mathscr{D}(\Omega)$ containing 0. Due to Theorem 9.3 (a), V is open in $\mathscr{D}(\Omega)$ if and only if $V \cap \mathscr{D}_K$ is open in $\mathscr{D}_K$ for every compact $K \Subset \Omega$. But this is precisely the content of (d), thereby completing the proof. ■

DEFINITION 9.5. A linear functional on $\mathscr{D}(\Omega)$ continuous with respect to the topology τ is called a *distribution*.

THEOREM 9.6. If Λ is a linear functional on $\mathscr{D}(\Omega)$, the following are equivalent:

(a) $\Lambda \in \mathscr{D}'(\Omega)$.

(b) To every compact $K \Subset \Omega$, corresponds a nonnegative integer N and a constant $C < \infty$ such that

$$|\Lambda\phi| \leqslant C\|\phi\|_N \qquad \forall\, \phi \in \mathscr{D}_K.$$

Proof. If $\Lambda \in \mathscr{D}'(\Omega)$, then the restriction of Λ to every $\mathscr{D}_K$ is continuous and so is bounded on some neighborhood of the origin, containing an open neighborhood of the form

$$\{\phi \in \mathscr{D}_K \colon \|\phi\|_N < \frac{1}{N}\},$$

whence (b) follows.

Conversely, suppose (b) holds. Then, as argued above, the restriction of Λ to every $\mathscr{D}_K$ is continuous, and due to the prededing theorem, Λ is continuous. ∎